

**North American Climate in CMIP5 Experiments.**

**Part II: Evaluation of Historical Simulations of Intra-Seasonal to Decadal**

**Variability**

Justin Sheffield, Suzana J. Camargo, Rong Fu, Qi Hu, Xianan Jiang, Nathaniel Johnson,  
Kristopher B. Karnauskas, Seon Tae Kim, Jim Kinter, Sanjiv Kumar, Baird  
Langenbrunner, Eric Maloney, Annarita Mariotti, Joyce E. Meyerson, J. David Neelin,  
Zaitao Pan, Alfredo Ruiz-Barradas, Richard Seager, Yolande L. Serra, De-Zheng Sun,  
Chunzai Wang, Shang-Ping Xie, Jin-Yi Yu, Tao Zhang, Ming Zhao

Justin Sheffield, Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ  
Baird Langenbrunner, Joyce E. Meyerson, J. David Neelin, Department of Atmospheric and Oceanic  
Sciences, University of California Los Angeles, CA  
Suzana J. Camargo, Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY  
Rong Fu, Jackson School of Geosciences, University of Texas at Austin, TX  
Qu Hu, School of Natural Resources and Department of Earth and Atmospheric Sciences, University of  
Nebraska-Lincoln, Lincoln, NE  
Xianan Jiang, Joint Institute for Regional Earth System Science and Engineering, University of California,  
Los Angeles, CA  
Kristopher B. Karnauskas, Woods Hole Oceanographic Institution, Woods Hole, MA  
Seon Tae Kim, CSIRO, Marine and Atmospheric Research, Aspendale, Victoria, Australia  
Sanjiv Kumar, Jim Kinter, Center for Ocean-Land-Atmosphere Studies, 4041 Powder Mill Road, Suite  
302, Calverton, MD  
Eric D. Maloney, Department of Atmospheric Science, Colorado State University, Fort Collins, CO  
Annarita Mariotti, National Oceanic and Atmospheric Administration, Office of Oceanic and Atmospheric  
Research (NOAA/OAR), Silver Spring, MD  
Zaitao Pan, Saint Louis University, St. Louis, MO  
Alfredo Ruiz-Barradas, Department of Atmospheric and Oceanic Science, University of Maryland, College  
Park, MD  
Richard Seager, Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY  
Yolande L. Serra, Department of Atmospheric Science, University of Arizona, Tucson, AZ  
De-Zheng Sun, Cooperative Institute for Environmental Studies/University of Colorado & NOAA/Earth  
System Research Laboratory, Boulder, CO  
Chunzai Wang, Physical Oceanography Division, NOAA Atlantic Oceanographic and Meteorological  
Laboratory, Miami, FL  
Jin-Yi Yu, Department of earth System Science, University of California, Irvine, CA  
Nathaniel Johnson, International Pacific Research Center, SOEST, University of Hawaii at Manoa,  
Honolulu, HI  
Shang-Ping Xie, International Pacific Research Center, and Department of Meteorology. SOEST,  
University of Hawaii at Manoa, Honolulu, HI  
Tao Zhang, Cooperative Institute for Environmental Studies/University of Colorado & NOAA/Earth  
System Research Laboratory, Boulder, CO  
Ming Zhao, NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NJ

43

44

Journal of Climate

45

Submitted July 30, 2012

46

Revised February 15, 2013

47

\*Corresponding author address: Justin Sheffield, Department of Civil and Environmental

48

Engineering, Princeton University, Princeton, NJ, 08540. Email: [justin@princeton.edu](mailto:justin@princeton.edu)

49

## 49    **Abstract**

50    This is the second part of a three-part paper on North American climate in CMIP5 that  
51    evaluates the 20<sup>th</sup> century simulations of intra-seasonal to multi-decadal variability and  
52    teleconnections with North American climate. Overall, the multi-model ensemble does  
53    reasonably well at reproducing observed variability in several aspects, but does less well  
54    at capturing observed teleconnections, with implications for future projections examined  
55    in part three of this paper. In terms of intra-seasonal variability, almost half of the models  
56    examined can reproduce observed variability in the eastern Pacific and most models  
57    capture the midsummer drought over Central America. The multi-model mean replicates  
58    the density of traveling tropical synoptic-scale disturbances but with large spread among  
59    the models. On the other hand, the coarse resolution of the models means that tropical  
60    cyclone frequencies are under predicted in the Atlantic and eastern North Pacific. The  
61    frequency and mean amplitude of ENSO are generally well reproduced, although  
62    teleconnections with North American climate are widely varying among models and only  
63    a few models can reproduce the east and central Pacific types of ENSO and connections  
64    with US winter temperatures. The models capture the spatial pattern of PDO variability  
65    and its influence on continental temperature and West coast precipitation, but less well  
66    for the wintertime precipitation. The spatial representation of the AMO is reasonable but  
67    the magnitude of SST anomalies and teleconnections are poorly reproduced. Multi-  
68    decadal trends such as the warming hole over the central-southeast US and precipitation  
69    increases are not replicated by the models, suggesting that observed changes are linked to  
70    natural variability.

71

## 1. Introduction

This is the second part of a three-part paper on the Climate Model Intercomparison Project phase 5 (CMIP5; Taylor et al., 2012) model simulations for North America. This second part evaluates the CMIP5 models in their ability to replicate the observed variability of North American continental and regional climate, and related climate processes. The first part (Sheffield et al., 2012) evaluates the representation of the climatology of continental and regional climate features. The third part (Maloney et al., 2012) describes the projected changes for the 21<sup>st</sup> century.

The CMIP5 provides an unprecedented collection of climate model output data for the assessment of future climate projections as well as evaluations of climate models for contemporary climate, the attribution of observed climate change and improved understanding of climate processes and feedbacks. As such, these data contribute to the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5), and other global, regional and national assessments.

The goal of this study is to provide a broad evaluation of CMIP5 models in their depiction of North American climate variability. It draws from individual work by investigators within the CMIP5 Task Force of the US National Oceanic and Atmospheric Administration (NOAA) Modeling Analysis and Prediction Program (MAPP) and is part of a Journal of Climate special collection on North America in CMIP5. We draw from individual papers within the special issue, which provide more detailed analysis that can be presented in this synthesis paper.

We begin in Section 2 by describing the CMIP5, providing an overview of the models analyzed, the historical simulations and the general methodology for evaluating



the models. Details of the main observational datasets to which the climate models are compared are also given in this section. The next 5 sections focus on different aspects of North American climate variability, organized by the time scale of the climate feature. Section 3 covers intraseasonal variability with focus on variability in the eastern Pacific Ocean and summer drought over the southern US and Central America. Atlantic and east Pacific tropical cyclone activity is evaluated in Section 4. Interannual climate variability is assessed in Section 5. Decadal variability and multi-decadal trends are assessed in Sections 6 and 7, respectively. Finally, the results are synthesized in Section 8.

## **2. CMIP5 Models and Simulations**

### *2.1. CMIP5 Models*

We use data from multiple model simulations of the “historical” scenario from the CMIP5 database. The CMIP5 experiments were carried out by 20 modeling groups representing more than 50 climate models with the aim of further understanding past and future climate change in key areas of uncertainty (Taylor et al., 2012). In particular, experiments have been focused on understanding model differences in clouds and carbon feedbacks, quantifying decadal climate predictability and why models give different answers when driven by the same forcings. The CMIP5 builds on the previous phase (CMIP3) experiments in several ways. Firstly a greater number of modeling centers and models have participated. Secondly, the models are more comprehensive in terms of the processes that they represent and are run at higher spatial resolution, therefore hopefully resulting in better skill in representing current climate conditions and reducing uncertainty in future projections. Table 1 provides an overview of the models used. The

specific models used vary for each individual analysis because of data availability at the time of this study, and so the model names are provided within the results section where appropriate.

## *2.2. Overview of Methods*

Data from the “historical” CMIP5 scenario are evaluated, which is a coupled atmosphere-ocean mode simulation that is forced by historical estimates of changes in atmospheric composition from natural and anthropogenic sources, volcanoes, greenhouse gases and aerosols, as well as changes in solar output and land cover. Historical scenario simulations were carried out for the period from the start of the industrial revolution to near present: 1850-2005. Our evaluations are generally carried out for the last 30 years of the simulations, depending on the type of analysis and the availability of observations. For some analyses the only, or best available, data are from satellite remote sensing which restricts the analysis to the satellite period, which is generally from 1979 onwards. In other cases the observational data are very uncertain for particular regions and time periods (for example, precipitation in high latitudes in the first half of the 20<sup>th</sup> century) and this is noted in the relevant sub-section. For other analyses, multiple observational datasets are available and are used to capture the uncertainty in the observations. The observational datasets are summarized in Table 2 and further details of the datasets and data processing are given in the relevant sub-sections and figure captions. Where the comparisons go beyond 2005 (e.g. 1979-2008), data from the model RCP8.5 future projection scenario simulation (as this is regarded as closest to the business as usual trajectory) are appended to the model historical time series. About half the models have

multiple ensemble members, but we select the first ensemble member for simplicity and discuss the variability in the results across the ensemble where appropriate.

### **3. Tropical Intraseasonal Variability**

#### *3.1. MJO-related variability over the eastern Pacific and adjoining regions*

It has been well documented that convection over the eastern Pacific (EPAC) ITCZ and neighboring areas is characterized by pronounced intraseasonal variability (ISV) during boreal summer (e.g., Knutson and Weickmann, 1987; Kayano and Kousky, 1999; Maloney and Hartmann, 2000a; Maloney and Esbensen, 2003, 2007; de Szoeke and Bretherton, 2005; Jiang and Waliser, 2008, 2009, 2011). ISV over the EPAC exerts broad impacts on regional weather and climate phenomena, including tropical cyclone activity over the EPAC and the Gulf of Mexico, the summertime gap wind near the Gulfs of Tehuantepec and Papagayo, the Caribbean Low-Level Jet and precipitation, the mid-summer drought over Central America and Mexico, and the North American monsoon (e.g., Magana et al., 1999; Maloney and Hartmann, 2000b; Maloney and Hartmann, 2000a; Maloney and Esbensen, 2003; Lorenz and Hartmann, 2006; Serra et al., 2010; Martin and Schumacher, 2010).

Here, model fidelity in representing ISV over the EPAC and Intra-America Sea (IAS) region is assessed by analyzing daily output of rainfall and 850hPa winds from eighteen CMIP5 models. Figure 1 displays a Taylor diagram for summer mean (May-September) precipitation from the CMIP5 models over the EPAC domain (150°W-80°W; 5°S-30°N) compared to the TMPA precipitation. While the two HadGEM models (HadGEM2-CC and HadGEM2-ES) display the highest pattern correlations ( $\sim 0.93$ ), the

MRI-CGCM3 show the smallest RMS due to its better skill in simulating the spatial standard deviations of summer mean rainfall over the EPAC. In addition, four models (MPI-ESM-LR, CSIRO-MK3-6-0, CanESM2, and CNRM-CM5) also exhibit relatively better pattern correlations than other models.

The leading ISV modes over the EPAC based on observed and simulated rainfall fields are identified using a complex empirical orthogonal function (CEOF) approach (Maloney et al., 2008). CEOF analyses were applied to 30-90-day band-pass filtered daily rainfall anomalies and the spatial amplitude and phase for the first CEOF mode (CEOF1) based on TMPA are illustrated in Figures 2a and 2b. A single ensemble member was used for each model for 1981-2005. The TMPA data are available for a shorter time period (13 years) but the sensitivity of the results to different sample sizes (based on data from a selected model) was found to be small. Similar to Maloney et al. (2008), the maximum amplitude of the observed rainfall CEOF1 occurs over the far eastern part of the EPAC. Figure 2b illustrates the pattern of spatial phase of observed rainfall CEOF1. In agreement with previous studies, the observed leading ISV mode associated with the CEOF1 largely exhibits an eastward propagation, while a northward component is also evident (e.g., Jiang and Waliser, 2008; Maloney et al., 2008; Jiang et al., 2011).

Next, the fidelity of the CMIP5 models in simulating the leading EPAC ISV mode is assessed by calculating pattern correlations of the simulated rainfall CEOF1 against observations. To increase sampling, spatial patterns of rainfall anomalies associated with the CEOF1 based on both observations and model simulations are derived at two quadratic phases by multiplying the CEOF1 amplitude by the Cosine and

Sine of spatial phase at each grid point, respectively. The pattern correlations are then calculated at both of these two quadratic phases. A final pattern correlation for a particular model is derived by averaging these two pattern correlation coefficients. Figure 2c illustrates pattern correlations in depicting the CEOF1 rainfall pattern for each model simulation versus domain averaged CEOF1 amplitude relative to observations, which provide measures of model performance of variability in space and time, respectively. A majority of the CMIP5 models tend to underestimate the amplitude of the leading EPAC ISV mode associated with the rainfall CEOF1, except CNRM-CM5, MIROC5, MPI-ESM-LR, and HadGEM2-CC and HadGEM2-ES. Among the eighteen models examined, eight models exhibit relative higher pattern correlations ( $> 0.75$ ).

The models with relative better skill in representing the leading EPAC ISV mode also largely exhibit better skill for summer mean rainfall (cf. Fig. 1 and Fig. 2c) and 850hPa wind patterns (not shown). A common feature among the more skillful models is the presence of westerly or very weak easterly mean low-level winds over the EPAC warm pool region, as in the observations. Most of the models with relatively lower skill exhibit a stronger easterly summer mean flow ( $> 4$  m/s). This suggests that realistic representation of the mean state could be crucial for improved simulations of the EPAC ISV, which is in agreement with a recent study by Rydbeck et al. (2012), and has also been discussed for MJO simulations over the western Pacific and Indian Ocean (e.g., Kim et al., 2009). One hypothesis is that a realistic mean state produces the correct sign of surface flux anomalies relative to intraseasonal precipitation, which helps to destabilize the local intraseasonal disturbance (e.g. Maloney and Esbensen, 2005). Extended analyses of the EPAC ISV in CMIP5 models is given in Jiang et al. (2012).

### 3.2. Mid-summer Drought over Central America

The rainy season in Central America and southern Mexico spans roughly May through October. For most of the region, the precipitation climatology features maxima in June and September and a period of reduced rainfall during July-August known as the midsummer drought (MSD; Portig et al., 1961, Magaña et al., 1999). The MSD is regular enough to be known colloquially and plays an important role in farming practices (Osgood et al., 2009). A previous assessment of CMIP3 model performance at simulating the MSD and future projections (Rauscher et al., 2008) suggested that many models are capable of simulating the MSD despite an overall dry bias, and that the MSD is projected to become stronger with an earlier onset. In this section, the CMIP5 performance at simulating summertime precipitation and the MSD is evaluated. We evaluate 23 CMIP5 models against the TMPA, GPCP and UNAM observational datasets. A simple algorithm for detecting and quantifying the climatological MSD is used that does not assume *a priori* which months are maxima and which months constitute the MSD (Karnauskas et al., 2012).

Figure 3 shows the observational and CMIP5 estimates of the MSD and highlights the large uncertainties in its spatial distribution among observational datasets. The CMIP5 MME does reasonably well at representing the essence of the MSD over much of the Inter-Americas region. The maximum strength of the MSD in the MME is found just offshore of El Salvador and represents a midsummer precipitation minimum that is ~2.5 mm/day less than the early- and late-summer peaks. Significant differences in the location and strength of the MSD between the various observational data sets preclude a

definitive evaluation of the CMIP5 MME, but it is clear that the strength of the MSD is underestimated in some regions, including along the Pacific coast of Central America, the western Caribbean, the major Caribbean islands and Florida. Figure 3 also shows the MME standard deviation and a histogram of the spatial correlations of individual models with the MME mean. The largest uncertainties are collocated with the regions of largest magnitude of the MSD indicating that much of the model disagreement is in the magnitude. Several models stand out as outliers in representing the spatial distribution of the MSD relative to the MME mean (Table 3), such as MIROC-ESM and MIROC-ESM-CHEM, whilst the Hadley Center models do particularly well.

#### **4. East Pacific and Atlantic tropical storm track and cyclone activity**

##### *4.1. Tropical Storm Track*

The density of traveling synoptic-scale disturbances across the Tropics, referred to in the literature as the tropical storm track (e.g. Thorncroft and Hodges 2001; Serra et al. 2008; Serra et al. 2010), is examined in this section. These systems serve as precursors to a majority of tropical storms and hurricanes in the Atlantic and eastern North Pacific and their frequency at 850 hPa over Africa and the eastern Atlantic has been shown to be positively correlated with Atlantic hurricane activity (Thorncroft and Hodges 2001). As global models better resolve these systems than tropical cyclones, they provide an advantage over direct tracking of tropical cyclones to assess model tropical storm activity (see Section 4.2). As in Serra et al. (2010), the tropical storm track density is calculated based on the method of Hodges (1995; 1999) using smoothed, 6-hourly, 850 hPa relative

256 vorticity. Only positive vorticity centers with a minimum threshold of  $0.5 \times 10^{-6} \text{ s}^{-1}$  that  
257 persist for at least 2 days and have tracks of at least 1000 km in length are included in the  
258 analysis. This method primarily identifies westward moving disturbances such as easterly  
259 waves (e.g. Serra et al., 2010), although more intense storms that could potentially reach  
260 hurricane intensity are not excluded. We analyze a single ensemble member from nine  
261 CMIP5 models and compared the track statistics to the ERA-Interim (Figure 4, top).  
262 These models were selected based on whether the 6-hourly pressure level data were  
263 available at the time of the analysis. Mean track strength, the mean of the smoothed 850  
264 hPa vorticity along the track, is also examined (Figure 4, bottom).

265       The multi-model mean track density is in good agreement with ERA-Interim,  
266 however significant differences are seen with the individual models. The most apparent  
267 discrepancies are with the BCC-CSM1-1, CanESM2 and CCSM4 models, which strongly  
268 overestimate activity across the East Pacific and suggest a more longitudinally oriented  
269 track (CanESM2 and CCSM4) shifted south from what is observed. BCC-CSM1-1,  
270 HadGEM2-ES and MIROC5 underestimate tracks in the West Atlantic, while GFDL-  
271 ESM2M underestimates tracks throughout the region except near  $130^\circ\text{W}$ . MPI-ESM-LR  
272 also underestimates tracks across the region as well as shifts their location southward.  
273 The track density maximum off the west coast of Mexico is best captured by HadGEM2-  
274 ES, while the overall smallest magnitude differences are seen with CNRM-CM5. The  
275 multi-model mean track strength maximum in the East Pacific lies along the west coast of  
276 Mexico similar to ERA-Interim, however it is broader in scale and of larger magnitude  
277 than the observations (Figure 4, bottom). On the other hand, the multi-model mean  
278 strength in the Gulf of Mexico and West Atlantic along the east coast of the US is



strongly underestimated compared to ERA-Interim. Unlike for track density, these biases are fairly consistent among the models, with the exception of BCC-CSM1.1, which strongly overestimates mean strength across the region..

To better understand the biases in mean track density and strength we examine the spatial correlations of 850 hPa and 500 hPa winds and heights, as well as track density and strength with the ERA-Interim reanalyses. While all nine models have relatively good spatial correlations in the wind components and heights at 500 hPa (not shown), there is a wide spread in performance at the 850 hPa level that corresponds reasonably well with the rankings for the combined track density and strength correlations (Table 3). In particular, the top two models for the combined 850 hPa wind and height correlations (CNRM-CM5 and HadGEM2-ES) are also among the highest ranked for the combined track density and strength correlations. On the other hand, CanESM2 has a high ranking in the combined 850 hPa index but is one of the poorer models with respect to track density and spatial correlations, suggesting that there are other important factors contributing to the track statistics than just the large-scale low-level heights and winds across the region.

#### *4.2. Tropical Cyclones in the North Atlantic and Eastern North Pacific*

It is well known since the 1970s that climate models are able to simulate tropical cyclone-like storms (e.g. Manabe et al., 1970; Bengtsson et al., 1982), which are generally formed at the scale of the model grid when conditions are unstable enough and other factors, such as vertical wind shear, are favorable. As the resolution of the climate models increases, the modeled storm characteristics become more realistic (e.g. Zhao et

al., 2009). Analysis of CMIP3 models showed that the tropical cyclone-like storms produced still had many biases common of low-resolution models (Walsh et al., 2010). Therefore, various dynamical and statistical techniques for downscaling tropical cyclone activity using only the CMIP3 large-scale variables have been employed (Emanuel et al., 2008; Knutson et al., 2008). Recent studies suggest that when forced by observed SSTs and sea-ice concentration, a global atmospheric model with a resolution ranging from 50 to 20km can simulate many aspects of tropical cyclone (TC)/hurricane frequency variability for the past few decades during which reliable observations are available (e.g., Oouchi et al. 2006; Bengtsson et al. 2007; Zhao et al. 2009). The success is not only a direct evaluation of model capability but also an indication of the dominant role of SST variability on TC/hurricane frequency variability. When assuming a persistence of SST anomalies, some of the models were also shown to exhibit significant skill in hurricane seasonal forecast (e.g., Zhao et al. 2010; Vecchi et al. 2011).

Tropical storms and cyclones in this study are identified using the tracking method of Camargo and Zebiak (2002), which uses low-level vorticity, surface winds, surface pressure, and atmospheric temperature, and considers only warm core storms. The method uses model-dependent (and resolution) thresholds and storms have to last at least two days. Only a subset of the tropical disturbances examined in the previous section will intensify enough to be identified by this tracking method and the percentage that this occurs will vary among different models. As will be shown, the CMIP5 standard models have trouble simulating the number of tropical cyclones, which can be attributed in part to their coarse resolution. Therefore, we also show results from the GFDL high-resolution model.

TC type structures were tracked in five models for 1950-2005. We compare with observations from best-track datasets of the National Hurricane Center (Figure 5). The number of TCs in all models is much lower than in observations, which is common to many low-resolution global climate models (e.g. Camargo et al., 2005; 2007). The HadGEM2-ES has the largest low bias and the MPI-ESM-LR model has the most realistic tracks in the Atlantic basin. The MRI-CGCM3 model tracks in the Atlantic are mostly in the subtropical region, with very few storms in the deep tropics. In contrast, in the eastern North Pacific the MRI-CGCM3 has storm activity too near the equator. In the eastern North Pacific, very few storms (in all models) have westward tracks. The models seem to have an easier time in producing storms that are in the northwestward direction parallel to the Central American coast.

Figure 6 shows the mean number of TCs per month for the North Atlantic and eastern North Pacific. In some cases, the models produce too many storms in the off-season, while all models produce too few storms in the peak season. The bottom panels show the spread of the number of storms per year, emphasizing the low number of storms per year in all models. The highest resolution model MRI-CGCM3 ( $1.1^\circ \times 1.1^\circ$ ) has the least bias relative to the observations and the highest bias is for the coarsest resolution model (GFDL-ESM2M,  $2.5^\circ \times 2.0^\circ$ ). However, resolution cannot explain the rankings for all models, with the HadGEM2-ES and MPI-ESM-LR models having relatively large and small biases, respectively, despite both having intermediate resolutions. The model dynamical core, convection scheme and their interactions are other factors that have been shown to be important (Camargo, 2013). Examination of variability across ensemble

members in producing tropical cyclones was carried out for five member runs of the MRI-CGCM3 model (not shown) but was much less than among different models.

Figure 7 shows results for the GFDL C180HIRAM model, which has a higher-resolution (~50km) than the standard coupled GFDL-CM3 model and differs in some aspects of the physics such as the convection scheme. The model was run for a CMIP5 time slice experiment forced by observed inter-annually and seasonally varying SSTs and sea-ice concentration from the HadISST data-set (Held et al., 2013). The tracking algorithm of Zhao et al. (2009) was used to identify TCs with near surface wind speed reaching hurricane intensity. The model reproduces the observed statistics with the ratio of observed to model variances of interannual variability in both the N. Atlantic and E. Pacific not statistically different from one, according to an F-test at the 95% significance level. Figure 7 also shows that the model captures the observed seasonal cycle in both the N. Atlantic and E. Pacific. The model can also reproduce the observed seasonal cycle in the N. Atlantic and E. Pacific as well as the observed year-to-year variation of annual hurricane counts and the decadal trend for both basins for this period (Zhao et al. 2009; Held et al. 2013). The quality of the model's present-day simulation increases our confidence in the future projections, although the uncertainty in the projections is dominated by uncertainty in projected changes in SST boundary conditions across the CMIP5 standard resolution models (Maloney et al., 2013).

## **5. Interannual to decadal variability**

### *5.1. El Niño-Southern Oscillation (ENSO)*

The ENSO is the most important driver of global climate variability on inter-annual time scales. It impacts many regions worldwide through climate teleconnections (Ropelewski and Halpert, 1987), which link the tropical Pacific to higher latitudes through shifts in mid-latitude weather patterns. The impact of ENSO on North American climate is felt most strongly in the wintertime, with El Niño events bringing warmer temperatures to much of the northern part of the continent and wetter conditions in the southern US and northern Mexico. La Niña events tend to bring drier weather to the southern US. Evaluation of the ability of CMIP5 models to simulate ENSO is carried out for several aspects of ENSO variability and for teleconnections with North American climate.

#### A. Evaluation of ENSO teleconnections

We examine how well the historical simulations of CMIP5 models reproduce the composite near-surface air temperature (SAT) and precipitation patterns over North America during El Niño and La Niña episodes. In both model and observed data, we define ENSO episodes similarly to the Climate Prediction Center (CPC). A monthly ENSO index is calculated from detrended and high-pass filtered SSTs over the Niño3.4 region ( $5^{\circ}\text{S} - 5^{\circ}\text{N}$ ,  $170^{\circ}\text{W} - 120^{\circ}\text{W}$ ) from ERSSTv3b observations and CMIP5 models. An El Niño (La Niña) episode is defined as any sequence of months where the three-month running mean Niño3.4 SST, is greater than  $0.5^{\circ}\text{C}$  (less than  $-0.5^{\circ}\text{C}$ ) for at least five consecutive three-month running seasons.

In observations approximately 90% of El Niño and 89% of La Niña episodes feature peak amplitudes in fall or winter. In the CMIP5 ensemble of the historical

simulations, however, only 68% of El Niño and 65% of La Niña episodes have peak amplitudes in fall or winter, although several of the models (CanESM2, CNRM-CM5, HadCM3 and NorESM1-M) do have fall/winter peak frequencies exceeding 80% for both El Niño and La Niña episodes. This finding suggests that CMIP5 models do not fully reproduce the phase-locking of ENSO to the seasonal cycle, a deficiency noted in CMIP3 models as well (Guilyardi et al. 2009). The following analysis focuses on those episodes that do peak in fall or winter. In the ensemble mean, the frequency of ENSO episodes and the mean peak amplitude are similar to observed values (not shown).

Because the dynamics of extratropical ENSO teleconnections are tied to upper tropospheric processes, and because these teleconnections are strongest during boreal winter, we examine how well CMIP5 models reproduce the DJF composite 300hPa geopotential height patterns in the NCEP/NCAR reanalysis. In addition, we attempt to identify what characteristics distinguish higher from lower performance models, where performance is based on the El Niño (La Niña) composites of all height fields for which the detrended Niño 3.4 SST anomaly is greater than 0.5°C (less than -0.5°C). The high performance models are defined as those with a pattern correlation that exceeds 0.6 and an RMS difference less than 13 m between the model and observed composites for both El Niño and La Niña (Figure 8). This subjective partitioning is used as a means of discerning general properties that distinguish higher from lower performance models. Overall, ten (eleven) models are characterized as high (low) performance based on these criteria.

Figure 9 shows the composites of 300 hPa geopotential height, SAT, precipitation, and tropical SST for El Niño. The corresponding composites for La Niña

(not shown) are quite similar but of opposite sign. The higher performance ensemble performs rather well in capturing the basic El Niño geopotential height, SAT, and precipitation teleconnections over the North Pacific and North America, with the exception being the failure to capture the negative precipitation anomaly in the Tennessee and Ohio Valleys. The lower performance ensemble features a much weaker teleconnection pattern and an Aleutian low anomaly that is shifted about 10° too far west. The composite El Niño SST anomalies (Figs. 2k,l), however, are quite similar.

To gain insight into possible reasons for the discrepancies between the higher and lower performance ensemble, Figure 10a shows composite differences in tropical precipitation. The higher performance ensemble exhibits much higher precipitation anomalies in the central and eastern equatorial Pacific Ocean, which suggests that the enhanced convection in these regions could help to explain the stronger and eastward shifted teleconnection pattern relative to the lower performance ensemble. This enhanced convection may be explained in part by stronger SST anomalies in the higher performance ensemble (Fig. 10b), but most of the large precipitation differences actually occur where the SST anomaly differences are quite small. Instead, a more significant difference appears to be the difference in SST climatology, as the lower performance ensemble exhibits climatological SSTs more than 1°C cooler than the high performance ensemble over the eastern Pacific cold tongue region (Fig. 10c). Indeed, the lower performance ensemble features a negative SST climatology bias of more than 1.5°C in the equatorial central Pacific (Fig. 10e), where the El Niño convection anomalies generally are strongest. The bias for the higher performance ensemble in this region (Fig. 10d) is much weaker. Thus, in the lower performance ensemble, the convection

anomalies in the eastern Pacific likely are too insensitive to ENSO SST anomalies because the climatological SSTs are too low. This finding suggests that simulation of ENSO teleconnections in some climate models might benefit from improving climatological SSTs rather than interannually varying ENSO SST anomalies. As discussed in Li and Xie (2012), tropical SST biases in CMIP models are linked to model errors in cloud cover and ocean dynamics, with equatorial cold tongue biases closely tied to errors in thermocline depth and upwelling.

#### B. East Pacific/Central Pacific ENSO and Teleconnections with US Winter Surface Air Temperature

It has been increasingly recognized that different types of ENSO occur in the tropical Pacific (e.g. Wang and Weisberg, 2000; Trenberth and Stepaniak, 2001; Larkin and Harrison, 2005; Yu and Kao, 2007; Ashok et al., 2007; Kao and Yu, 2009; Kug et al., 2009). Two particular types that have been emphasized are the Eastern-Pacific (EP) type that produces SST anomalies near the South America coast and the Central-Pacific (CP) type that produces anomalies near the international dateline. While the EP ENSO is the conventional type of ENSO, the CP ENSO has gradually increased its occurrence during the past few decades (e.g. Lee and McPhaden, 2010). Recent observational studies have indicated that the impacts produced by these two types of ENSO on North American climate can be different (e.g., Mo 2011; Yu et al. 2012; Yu and Zhou, 2013). Here the ENSO teleconnection over the US simulated in the CMIP5 models are further examined according to the ENSO type. Following Kao and Yu (2009) and Yu and Kim (2010), a regression-EOF analysis is used to identify the CP and EP types from monthly SSTs. The



SST anomalies regressed with the Niño1+2 SST index were removed before the EOF analysis was applied to obtain the spatial pattern of the CP ENSO. Similarly, we subtracted the SST anomalies regressed with the Niño4 SST index before the EOF analysis was applied to identify the leading structure of the EP ENSO. The principal components of the leading EOF modes represent the ENSO strengths and are defined as the CP ENSO index and the EP ENSO index. The observed winter (DJF) SAT anomalies regressed to these two indices are different over the US (Figure 11a,b) with a warm northeast to cold southwest pattern for the EP El Niño and a warm northwest to cold southeast pattern for the CP El Niño. Adding these two impact patterns together results in a pattern that resembles the well-known warm-north, cold-south pattern of El Niño impact.. The robustness of these two different impact patterns has been examined in Yu et al. (2012) using numerical model experiments and case studies. They showed that impact patterns similar to those shown in Figure 11 can be reproduced in two ensemble AGCM experiments forced separately by the EP and CP ENSO SST anomalies (see their Fig. 1). The regressed impact patterns can also be identified in US winter temperature anomalies during the four strongest EP El Niño events (i.e., 1997-98, 1982-83, 1972-73, and 1986-87) and three of the four strongest CP El Niño events (i.e., 2009-10, 1957-58, and 2002-2003).

We repeated the EOF and regression analyses to evaluate how well the CMIP5 models reproduce the different US impacts to the two types of ENSO. The regressed winter SAT anomaly patterns calculated from 22 CMIP5 models are shown in Figure 11. The observed patterns are well simulated by some models, such as the MIROC5 and MRI-CGCM3 for the EP ENSO and the NorESM1-M and HadGCM2-ES for the CP

ENSO. However, some models show an impact pattern that is almost opposite to that observed, such as HadCM3 for the CP ENSO and INMCM4 for the EP ENSO. To quantify how well the impact patterns are simulated, pattern correlation coefficients were calculated between the model regressed patterns and the NCEP regressed patterns. As shown in Figure 12a, there is a cluster of eleven CMIP5 models (CSIRO-Mk3-6-0, GFDL-CM3, GFDL-ESM2G, GFDL-ESM2M, HadGEM2-CC, HadGEM2-ES, IPSL-CM5-MR, MIROC5, MPI-ESM-LR, MPI-ESM-P, NorESM1-M) that have higher pattern correlation coefficients for both the EP ENSO and the CP ENSO than the rest of the models. This group of the CMIP5 models is considered as the models whose regressed US winter temperature patterns are close to the observed patterns for the two types of ENSO. We also examine in Figure 12b the intensities of the simulated EP and CP ENSO events, which are determined using an EOF-regression method (Yu and Kim, 2010; Kim and Yu, 2012). Models with realistically strong events are identified using the lower limit of the 95% significance interval of the observed intensities (using an F-test) as the criteria (0.78°C for EP and 0.51°C for CP). Based on these criteria, ten of the 22 models simulate both EP and CP ENSO events with realistically strong intensities. Interestingly, nine of these models are also among the eleven models that realistically produce US winter temperature patterns for the two types of ENSO. Therefore, at least nine out of 22 models can more realistically produce the two types of ENSO with higher intensities and their different impacts on US winter temperatures: GFDL-CM3, GFDL-ESM2G, GFDL-ESM2M, HadGEM2-CC, HadGEM2-ES, MIROC5, MPI-ESM-LR, MPI-ESM-P, and NorESM1-M).

### C. ENSO warm/cold events asymmetry

ENSO asymmetry refers to the fact that the two phases of ENSO are not mirror images of each other (Burgers and Stephenson, 1999). The asymmetry shows up in both the surface and subsurface fields (Rodgers et al., 2004; Schopf and Burgman, 2006; Sun and Zhang, 2006; Zhang et al., 2009). Causes for such an asymmetry are not yet clearly understood, but accumulating evidence suggests that it is likely a consequence of nonlinearity of ocean dynamics (Jin et al., 2003; Sun 2010, Liang et al., 2012). Asymmetry is also linked to the time-mean effect of ENSO (Sun and Zhang, 2006; Schopf and Burgman, 2006; Sun, 2010, Liang et al. 2012). Understanding the causes and consequences of ENSO asymmetry may hold the key to understanding decadal variability in the tropics and beyond (Rodgers et al., 2004; Sun and Yu, 2009, Liang et al., 2012). Figure 13 shows the sum of the SST anomalies between the warm and cold phases of ENSO from HadISST observations and CMIP5 models. The threshold value used for defining the warm and cold phase anomalies is set as +0.5 °C and -0.5 °C respectively. This sum has also been called the SST anomaly residual and has been a common measure of the ENSO asymmetry in the SST field. All models underestimate the observed positive SST residual (and therefore the asymmetry) over the eastern Pacific. Measured by the skewness of Niño3 SST anomalies (which is a more rigorous measure of asymmetry), all the models also underestimate the observed ENSO asymmetry (Figure 14). The figure also shows that the stronger variability of ENSO (measured by variance) does not guarantee a stronger asymmetry in ENSO (measured by skewness).

ENSO asymmetry remains a common bias in climate models that has continued since CMIP3 (van Oldenborgh et al. 2005) with implications for simulating tropical

decadal variability. The causes are of current debate, but recent results indicate that it is related to the mean state and the excessive cold tongue in the models (De-Zheng Sun, personal communication, 2013), which was also noted in CMIP3 models (Sun et al., 2012), although there is evidence that the mean state could in turn be determined by the statistics of ENSO via non-linearities in the system (Sun and Zhang 2006, Sun 2010, Liang et al. 2012, Sun et al. 2012, Ogata et al. 2012). On other hand, both the bias in the mean state and the bias in the asymmetry may be a consequence of a more fundamental reason: a weak thermal forcing relative to the dissipation (Sun, 2000; Liang et al. 2012). Together, these results suggest that the coupled tropical system in the models is in a different dynamical regime to reality (Sun and Bryan, 2010), also noted in terms of the elevation of the variance of ENSO over the past 50 years that is not represented by the models (Sun, 2010).

## *5.2. Persistent droughts and wet spells over Great Plains and the southern-tier states*

Persistent dry and wet summers are features of the US Great Plains and southern US. We evaluate how the CMIP5 models describe the processes that cause such persistent anomalies in terms of low-level circulation and moisture flux anomalies by comparing with the NCEP-NCAR reanalysis. This complements the evaluations of the average seasonal circulation in the region, such as the low-level southerly jet as shown in part 1 of this paper (Sheffield et al., 2013). Persistent wet and dry summers are defined by JJA precipitation anomalies averaged over the Great Plains region from 90°-105°W and 30°-50°N during 1971-2000. Wet (dry) summers are identified as having normalized JJA precipitation larger (smaller) than 0.6 (-0.6) standard deviation. The reanalysis data

identify 8 wet and 7 dry summers in 1971-2000, and the models identify between 7 and 12 wet or dry events, depending on the model. We show the composites of vertically integrated moisture from the surface to top of the troposphere, the 850hPa geopotential height, and near surface winds at 925hPa for the wet and dry summers and their differences for the reanalysis (Figure 15) and for a single model, CCSM4, as an example (Figure 16).

Comparison of the two figures indicates some similarities but also very different processes causing the persistent wet or dry summers. The integrated moisture fluxes in both datasets indicate high moisture in an averaged cyclonic rotation in the troposphere in persistent wet summers (Figs. 15a and 16a) but anticyclonic rotation in dry summers (Figs. 15b and 16b) in the Great Plains. However, the sources of the moisture and the low-level dynamic structure are quite different. For the reanalysis, the convergence of moisture in the central Great Plains during wet summers results from southerly flow anomalies in the enhanced subtropical high pressure system in the North Atlantic and northerly flow anomalies in low pressure anomalies centered in the Midwest (Fig. 15d). These anomalies suggest a frontal system along the depression from the Midwest to the Southwest. A nearly reversed pattern of flow anomalies is shown during the dry summers (Figs. 15e and 15f). The model simulations show a different pattern of flow anomalies (Figs. 16d and e). In wet summers, the moisture is primarily from the east along the easterly and southeasterly quadrants of a high pressure anomaly center in the Great Lakes areas, instead of from the south as in the reanalysis result (Figs. 16a vs. 15a). In dry summers, the model shows dry flows from the Mexican plateau off the Sierra Madre Oriental in Mexico. These contrasts are shown in Fig. 16f. The other CMIP5 models also

simulate different tropospheric circulation patterns from those in the reanalysis for both wet and dry summers in the Great Plains.

Although the integrated moisture fluxes in the models resemble those in the reanalysis estimates in wet and dry summers, the sources of moisture differ considerably, suggesting that the models are not correctly representing the mechanisms that force variability in the Great Plains. Controls on summertime Great Plains precipitation have been found to depend strongly on moisture transport from the Gulf of Mexico via the Great Plains low level jet (GPLLJ; e.g., Ruiz-Barradas and Nigam, 2006; Cook et al., 2008; Weaver and Nigam, 2008) whose variability in turn may be related to remote SST forcing in the Pacific (e.g. Schubert et al., 2004; Ruiz-Barradas and Nigam, 2010, McCabe et al., 2008) and Atlantic (e.g. Enfield et al., 2001, Sutton and Hodson, 2005; McCabe et al., 2008) with contrasting anomalies in each basin associated with extreme conditions in the Great Plains (e.g., Hoerling and Kumar 2003; Schubert et al. 2009). Some of the models have shown improvement, compared to the CMIP3 models, in simulating the GPLLJ and the seasonal transitions (see Sheffield et al. 2013), a result largely attributable to the higher spatial resolution of CMIP5 models, but most models struggle to represent observed teleconnections between precipitation and Atlantic SSTs (see section 6). Even so, the transport of moisture transport is not the whole story and local dynamic processes (e.g. Veres and Hu, 2013), and land-atmosphere feedbacks (Ruiz-Barradas and Nigam, 2006), are important to initiate and further organize regional circulations that can transform the moisture into precipitation. Notably, previous studies focused on climate models find that they tend to over-estimate the role of recycled

precipitation over advected moisture (e.g. Ruiz-Barradas and Nigam, 2006) for the Great Plains with implications for the modeled precipitation variability.

## **6. Decadal Variability**

### *6.1. Pacific Decadal Oscillation (PDO) and its influence on North American climate*

On interdecadal timescales, variability in the tropical and extratropical North Pacific, particularly that of the Pacific Decadal Oscillation (PDO), has significant physical and ecological impacts over North America (Mantua et al. 1997, Higgins et al. 2000, Meehl et al. 2012). We examine the PDO and its relationships with N. American temperature and precipitation for 21 CMIP5 models. We define the PDO as the leading empirical orthogonal function of extended winter (November-April) monthly mean sea surface temperature (SST) anomalies in the North Pacific poleward of 20°N (Zhang et al., 1997; Mantua et al., 1997) for 1900-1993, and subtract the monthly global mean SST. We then calculate the PDO index by projecting monthly North Pacific SST anomalies onto the PDO pattern for all available months and then standardizing the resulting time series. Figure 17 illustrates the PDO patterns in both observations and the CMIP5 ensemble (see Table 5 for a list of models) obtained by regressing the unfiltered monthly SST anomalies onto the PDO index for all calendar months. As in the CMIP3 models (Oshima and Tanimoto, 2009; Furtado et al., 2011), the CMIP5 models reproduce the basic PDO horseshoe SST pattern. The most notable difference is the westward shift of the North Pacific center of action in models with respect to observations (Fig. 17c). The

regions with the largest differences also correspond with regions of relatively high inter-model variability (Fig. 17d).

For each set of seasonal temperature and precipitation regressions, we calculate the centered pattern correlations and RMS differences between the observed and CMIP5 model regressions (Table 5). Despite fairly low pattern correlations in many cases, for most models and most seasons, the differences in the regression patterns are not statistically significant. This may be due to a combination of small effective sample size, large uncertainty in the regression coefficients, a relatively modest impact of the PDO on seasonal SAT and precipitation, and the ability of the models to capture the general PDO behavior during the winter and spring when the PDO impacts are strongest. In particular, the full ensemble performs well in capturing the winter and spring PDO SAT patterns, but substantial differences in the precipitation regressions are evident, particularly in spring.

Figure 18 shows the DJF SAT and precipitation regressions in observations and the CMIP5 ensemble. The CMIP5 models do rather well in capturing the PDO influence on North American SAT, with positive (negative) SAT anomalies in northwest (southeast) North America during the positive phase of the PDO. Almost all local differences in the regression coefficients are not statistically significant. In contrast, the CMIP5 models perform somewhat poorly in reproducing the precipitation patterns over large parts of North America, although for high latitudes the observations are based on very sparse station data, and especially before the 1950s (Zhang et al., 2000). Both observations (Fig. 18b) and CMIP5 ensemble (Fig. 18d) produce a tripole pattern of precipitation anomalies over the west coast of North America. Large differences,



however, are found in eastern North America. In observations, the positive phase of the PDO is associated with reduced wintertime precipitation in the Tennessee and Ohio Valleys, northeastern U.S., and Southeastern Canada (Fig. 18b), but the CMIP5 ensemble fails to discern this influence (Fig. 18d, f). Though of smaller magnitude, significant differences also occur in central North America (Fig. 18f). In spring (MAM) the largest differences in the precipitation regressions occur along the coast of British Columbia, where observed regressions indicate positive anomalies but the CMIP5 ensemble produces a pronounced negative anomaly (not shown). Both observations and the CMIP5 ensemble reproduce positive precipitation anomalies along the west coast and central plains of the U.S.

## *6.2. Atlantic Multidecadal Oscillation (AMO)*

The Atlantic Multidecadal Oscillation (AMO) is an important mode of multidecadal climate variability manifesting in North Atlantic SSTs (e.g., Kerr 2000; Enfield et al. 2001). The AMO has significant regional and global climate associations, such as northeast Brazilian and Sahel rainfall (e.g., Folland et al. 1986; Rowell et al. 1995; Wang et al. 2012), hurricane activity in the North Atlantic and the eastern North Pacific (Goldenberg et al. 2001; Wang and Lee 2009), and North American and European summer climate (Enfield et al. 2001; McCabe et al. 2004; Sutton and Hodson 2005). In spite of its importance, the mechanism of the AMO is still unclear. Several studies have indicated the role of variations in the Atlantic meridional overturning circulation (AMOC) and associated heat transport fluctuations (Delworth and Mann 2000; Knight et al. 2005). Some modeling studies indicate that solar variability and/or volcanoes are

important (Hansen et al. 2005; Ottera et al. 2010), or that aerosols can be a primary driver (Booth et al. 2012). A recent observational study shows that a positive feedback between SSTs and dust aerosols in the North Atlantic via Sahel rainfall variability may be a mechanism (Wang et al. 2012).

The AMO index is defined as the detrended North Atlantic SST during the Atlantic hurricane season of June to November (JJASON) from the equator to 60°N, 75°W-5°W with the 11-year running mean (e.g., Enfield et al. 2001; Knight et al. 2005). As shown in Figure 19a, the individual models show highly varying amplitudes and phases, with a large spread across models. This is to be expected given that the AMO is partly of internal origin. However, the potential influence of external forcings implies that the models may simulate some of the variation. All models show the warming in the last two decades when anthropogenic warming becomes influential. The MME mean tends to follow the main variations in the earlier part of the record, albeit subdued because of averaging across models, but fails to show the warm period during 1926-1965. Compared to the CMIP3 results (Medhaug and Furevik, 2011), the CMIP5 simulation of the AMO has generally improved, particularly after 1960. This may be due to higher resolution, improved parameterizations and the addition of time evolving land cover. Results for individual models (Table 6) indicate that the standard deviations are comparable to, or slightly weaker than, the observations with typical amplitudes ranging from 0.09 to 0.19°C as compared to about 0.18°C in the observations, which is an improvement from CMIP3 models (Ting et al., 2009).

The lagged autocorrelation of the AMO index for lags zero to 35 years (Fig. 19b) shows that the models generally represent the quasi-periodic nature of the observed

AMO, with the peak oscillation at 30-35 years in the observation, but generally shorter for the models. The persistence in the AMO index as defined as the maximum time lag when the autocorrelation first crosses the significance line at the 90% level, varies from 5 to 25 years in the models, implying the potential for predicting future SSTs (Corti et al., 2012; Kim et al., 2012). However, for most models the persistence is shorter (~12 years), which is nevertheless an improvement over CMIP3 models which have an average persistence of about 5 years (Medhaug and Furevik 2011).

The ability of the models to represent the AMO and its impact on precipitation over North America is evaluated by regressing the AMO index on regional seasonal precipitation and SSTs for 1901-1999. The results are shown for autumn in Figure 20 and shown in more detail in Kavvada et al. (2013). The SST signature of the AMO is stronger in autumn than in summer and this is reflected in its impact on central US precipitation in observations (not shown). In both seasons the SST anomalies reach a maximum over the mid Atlantic, over the sub-polar gyre region. The warm phase of the AMO induces drying conditions over the central US and wet conditions over Florida and the US northeast in both seasons, but with more intensity in autumn. However, there are seasonally contrasting conditions along the Gulf of Mexico states where decreased precipitation occurs in summer but increased precipitation occurs in autumn.

In general the models do not capture the SST seasonality of the AMO well. The simulated SST anomalies are generally larger in summer than in autumn in the majority of the models (not shown). While all models tend to place the maximum SST anomalies over the mid Atlantic Ocean, they do not replicate the observed maximum south of Greenland and its spatial structure. For example, CCSM4, GFDL-ESM and MIROC5

emphasize anomalies over the Norwegian Sea; and GFDL-ESM, GISS-E2-R and INMCM4 do not show a signal over the tropical Atlantic. The spatial correlation of the anomalies (Table 7) shows higher correlations for HadGEM2-ES and GISS-E2-R, although visually there are large discrepancies in the spatial patterns.

The precipitation impact of the AMO is a bigger challenge for the models (see Table 7 for individual model spatial correlations for precipitation) and they generally fail to represent the drier conditions over the central US and the wet conditions along the coastal south Atlantic US states and southern Mexico. The initial drying over the south central US in summer is shown by a few models (BCC-CSM1.1, HadGEM2-ES, IPSL-CM5A-LR and MRI-CGCM3) but the intensification of the drying into the autumn is not replicated by most of the models. The wet conditions over the south Atlantic US states in the autumn are captured by a few models, but to varying degrees of agreement and some models show regressions of the opposite sign (e.g. GISS-E2-R and HadGEM2-ES) and despite their high SST correlations. The increased precipitation over southern Mexico in autumn is shown only by a handful of models (e.g. BCC-CSM1.1, CSIRO-MK3.6, IPSL-CM5A-LR and NorESM1-M).

Numerous studies have shown the importance of the AMO in generating precipitation variability over the region (e.g. Enfield et al 2001, Sutton and Hodson 2005, Wang et al. 2006, Schubert et al. 2009, Nigam et al. 2011), with a key role played by the lower level circulation which modulates the Great Plains low-level jet and the convergence/divergence of moisture fluxes (see section 5.2). Thus, given the differences in the model simulated structure of the AMO SST footprint, their poor performance in the

simulation of the hydroclimate impact over the central US is not surprising - a situation which has not shown improvement since CMIP3 (Ruiz-Barradas et al. 2013).

## **7. Multidecadal Trends**

### *7.1. Trends in Temperature and the ‘warming hole’ over the southeastern US*

A unique of feature of US temperature change during the 20<sup>th</sup> century is the so-called “warming hole (WH)” observed in the southeastern US (Pan et al, 2004). While the globe has warmed over the 20<sup>th</sup> century, the WH region experienced cooling, especially in summer during the latter half of the century. Studies have attributed the mechanisms for this abnormal cooling (lack of warming) trend to large-scale decadal oscillations such as PDO and AMO (Robinson et al., 2002, Kunkel et al., 2006, Wang et al., 2009; Weaver, 2012; Meehl et al., 2012) and to regional scale hydrological processes (Pan et al., 2004) and land surface interactions (Liang et at. 2007). Portmann et al. (2009) speculated that secondary organic aerosols during the growing season could contribute to the cooling in the WH region, while Christidis et al. (2010) emphasized the role of internal climate variability.

We evaluate whether the CMIP5 models show the warming hole as a forced response in Figure 21, which shows the annual and seasonal trends, in near surface air temperature from the observation and the CMIP5 multi-model mean from 17 models (see Figure 21 caption). Model and observation data are re-gridded to a common resolution  $2.5^{\circ} \times 2.5^{\circ}$  using area averaging. Trends are calculated for the 1930-2004 period using the Theil-Sen approach (Theil, 1950; Sen, 1968). The choice of 1930-2004 gives a

prominent WH signal in the observations starting from the warmest decade following the Dust Bowl drought. Only one ensemble member from each model is included in the analysis as ensemble members from the same model show similar spatial patterns of long-term trends (Kumar et al., 2013). The MME mean neither shows a cooling trend in the eastern US, nor lesser warming relative to the western US. This indicates that, similar to CMIP3 (Kunkel et al. 2006) simulations, the CMIP5 simulations do not show the WH as a forced response signal.

Figure 22 shows the temporal evolution of 30-year moving window annual temperature trends over the eastern US in the observational data and CMIP5 simulations, and relative to the western US. The multi-decadal persistence of the WH is clearly visible in the observational data i.e. most negative temperature trends are clustered between 1925 and 1955. The 95% model spread range brackets the observed multi-decadal variability in the eastern US temperature trends and approximately 40% of the 95% model spread range is negative. The multi-model median captures the overall tendency of positive and negative trend evolution ( $r^2 = 0.58$ ). Pan et al. (2013) found that 19 out of 100 CMIP5 historical ‘*all forcings*’ simulations showed negative temperature trends in the southeast USA; whereas simulations based on greenhouse gas emissions forcing only showed a strong warming in the central US. These results suggest that there is some fidelity with observations via external forcings, but natural climate variability plays a major role. Kumar et al. (2012) found that the 30 year running temperature trend variability in the eastern US is significantly correlated ( $r^2 = 0.76$ ) with the AMO and models that have relatively higher skill in AMO simulations also have a higher chance of reproducing the WH in the eastern United States. There is essentially no skill in the model’s

representation of the difference in trends between the eastern and western US running trends (Fig. 22b).

## *7.2. Trends in Diurnal Temperature Range (DTR)*

Observed warming during the day and night has been asymmetric, with nocturnal minimum surface air temperature (Tmin) rising about twice as fast than daytime maximum temperature (Tmax) during the second half of 20<sup>th</sup> century, mostly during 1950-1980 (Vose et al., 2005). Changes in cloud cover, atmospheric water vapor, soil moisture and other factors, account for 25-50% of the DTR reduction (Dai et al., 1999). Cloud cover, soil moisture, precipitation, and atmospheric/oceanic teleconnections account for up to 80% of regional variance over 1901-2002. Over the U.S., cloud cover alone accounts for up to 63% of regional annual DTR variability (Lauritsen and Rogers, 2012). During 1950-2004, summer Tmax and Tmin over North America increased 0.07 and 0.12 °C, respectively, resulting in a -0.05 °C decrease in DTR (Voss et al., 2005). A similar decrease (-0.06 °C) occurred in winter. Over the WH region, summer Tmax decreased sharply (-0.13 °C) while Tmin increased slightly (0.05 °C), yielding a DTR decrease of 0.18 °C. Winter DTR also decreased by 0.13 °C.

Figure 23 shows a comparison of DTR magnitude and the linear trend in DTR from 17 models against the CRU TS3.1 observational dataset. The observed mean DTR (Tmax-Tmin) is characterized by high values over the western high mountainous regions in summer and low values in high latitudes (Fig. 23a). The MME mean simulates this general pattern with under-estimation in the mountains. The observed DTR trend is predominantly negative in the U.S and Mexico and largely positive in Canada in both

seasons (Fig. 23b). The largest decreasing DTR trend up to 0.2°C per decade is over the southeastern U.S. warming hole region in summer. The model DTR trend is poorly reproduced, missing the extensive negative trend over the southeastern region where models simulated increasing DTR trend (Fig. 23b, right panels). The pattern correlation between the observed and simulated DTR is from 0.40 to 0.82, with a mean of 0.67 for the 17 models, but the correlation of DTR trend is much lower, ranging from 0.19 to -0.26 (mean = 0.03). The model skill in simulating DTR trends does not appear to have improved from CMIP3 (Zhou et al., 2009) and earlier model comparisons (e.g. Branagnza et al., 2004), however, the role of anthropogenic forcings appears to be essential in producing a decline in DTR (Zhou et al., 2009), even if it is underestimated.

### *7.3. Trends in Precipitation*

Precipitation has generally increased over North America in the last half of the 20<sup>th</sup> century (Karl and Knight, 1998; Zhang et al., 2000). Trends in precipitation are positively correlated with streamflow trends, thereby affecting water resource availability and flood potential (Lettenmaier et al. 1994; McCabe and Wolock 2002, Kumar et al. 2009). Figure 24 shows the multi-model ensemble average precipitation trend for 1930-2004 from 17 models against the CRU observations. The multi-model average weakly captures the wetting trend in North America, particularly at higher latitudes. Note that the precipitation gauge density before the 1950s was very low, especially in high latitudes, and the observational trends are very uncertain, especially in high-latitudes, at least for the first part of the time period. However, the MME mean fails to capture the trend magnitude, for example, the higher wetting trend (>20 mm/decade) in the eastern US.



Figure 25 (a) and (b) show the 30 year running trend during the 20th century in the eastern and western US, respectively. The 95% model spread brackets the observed precipitation trend magnitude in both regions. The higher wetting trend in the observations has slowed down in the last decade in the eastern US. The muted magnitude of the trend in Figure 24 seems to be a result of low signal to noise ratio (the multi-model median line hovers around the zero line in Figure 25), rather than a robust feature of CMIP5 climate models. Some individual models capture very well the observed trend magnitude. Drying in Mexico is a dominant but incorrect feature in the CMIP5 simulations, which is symptomatic of CMIP3 models also (IPCC, 2007) and is likely driven by the inadequate connection between increasing precipitation and global SST warming, at least for summer, in the majority of models as shown by Fu et al. (2012) for the southern US.

## **8. Discussion and Conclusions**

This study has evaluated the simulated variability from the CMIP5 multi-model ensemble at intra-seasonal to multi-decadal time scales for North America and adjoining seas. The results show a mixture of performance, with some aspects of climate variability well reproduced (e.g. the spatial footprint of the PDO and its teleconnections), others reproduced well by some models but not others (e.g. ISV in the tropical Pacific; ENSO teleconnections and types) and others poorly by most models (e.g. tropical cyclone frequency; ENSO asymmetry; teleconnections with the AMO; long-term trends in DTR and precipitation). No one model stands out as better than the others, but certain models do perform much better for certain features. For example, the Hadley Center models do

well for the Central America mid-summer drought and the SST footprint of the AMO; the MRI-CGCM3 model does relatively well for intra-seasonal and inter-annual variability in the tropical Pacific and for tropical cyclone counts. In general, higher-resolution models do better for features such as tropical cyclones, but this does not appear to be a dominant factor for other aspects of climate variability. Furthermore, no model stands out as being particularly unskillful, bolstering the argument to consider all models irrespective of performance to encompass the uncertainties (Knutti, 2010). In fact, the range of processes and metrics analyzed is a key advantage of this study, because skill in one aspect does not necessarily mean good performance in another. For example, the NorESM1-1 model does very well at representing the two types of ENSO and its teleconnections, but does poorly at representing ENSO asymmetry. As a consequence, an overall ranking of models, albeit seemingly attractive, is difficult given the challenges in quantitatively comparing performance across different types of analysis, as well as the logistical challenges of sampling the same set of models across all analyses.

For the climate features and models analyzed here, there does not appear to be a great deal of improvement since CMIP3. For example, CMIP5 models still cannot capture the seasonal timing of ENSO events, that tend to peak in the fall and winter, and the spurious drying signal in the southern US and Mexico continues from CMIP3. However, some features continue to be well simulated, such as the SST pattern of the PDO, and features related to spatial resolution are likely to have improved, such as the representation of TCs. Overall, the models are less able to capture observed variability and long-term trends than they are the mean climate state as evaluated in the first part of this paper (Sheffield et al., 2013), although this may be a result of model tuning to

observations (Räisänen, 2007). This is understandable for decadal to multi-decadal variability which is dependent on the models' internal variability or the sensitivity to external forcing, for which the observations can be very uncertain. Some of the biases in variability, however, appear to be related to problems in simulating the mean state, and there are encouraging signs that improvements in the models, or at least the understanding of the sources of errors, can be made (e.g. biases in the depiction of the mean state of tropical Pacific may be linked to biases in the ISV, the lack of asymmetry in ENSO phases and to teleconnections with North American climate).

The results have implications for the interpretation and robustness of the model projected future changes. The third part of this paper (Maloney et al., 2013) evaluates the model projections for a subset of the features analyzed in the first part of this paper (Sheffield et al., 2013) and this second part. As noted in the first part, the accurate simulation of historic climate features is not sufficient for credible projections, although the depiction of large-scale climate features is necessary. Several studies of future projections show only small differences between models that do better at replicating observations and those that do worse (e.g. Brekke et al., 2008; Knutti et al., 2010) whilst others have found relationships between model performance and future projections that can be related to physical processes (e.g. Hall and Qu, 2006; Boe et al, 2009). However, these types of studies are generally specific to certain climate features that do not necessarily provide confidence or pessimism in model skill in a broader sense.

The adequate depiction of the variability is nevertheless necessary because this is generally associated with the more extreme aspects of climate that impose the largest impacts. Furthermore, the depiction of the teleconnections associated with large-scale

variability is especially important because the impacts of potential changes in the variability of, say, ENSO (Van Oldenborgh et al., 2005; Muller and Roeckner, 2008) are subject to uncertainties in the representation of teleconnections (Maloney et al., 2013). Model variability can also have a large impact on future changes because the signal to noise ratio can be highly dependent on the model's natural variability resulting in misleading assessments of future changes and uncertainties across models (Tebaldi et al., 2011). The ability of the models to reproduce the observed trends may be a better indicator of model reliability than depiction of the mean climate or even its variability, because this indicates the model's sensitivity to an external forcing that may continue into the future, such as greenhouse gas concentrations. The problem here is that the trend analyzed is subject to uncertainties in the observations, the complications of natural variability in the real world and models, and uncertainties in feedbacks and how they may change in the future (Räisänen, 2007; Knutti, 2010). The generally poor ability of the models to reproduce the trends in precipitation, DTR and some features of regional temperature shown here are indicative of this.

**Acknowledgements.** We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. The authors acknowledge the support of NOAA Climate Program Office

919 Modeling, Analysis, Predictions and Projections (MAPP) Program as part of the CMIP5  
920 Task Force.  
921

## References

- Adler, R. F., G. J. Huffman, A. Chang, R. Ferraro, P. Xie, J. Janowiak, B. Rudolf, U. Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, and P. Arkin, 2003: The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present). *J. Hydrometeor.*, **4**, 1147–1167.
- Aiyyer, A. R., and C. Thorncroft, 2006: Climatology of vertical wind shear over the tropical Atlantic. *J. Clim.*, **19**, 2969–2983.
- Aiyyer, A. and J. Molinari, 2008: MJO and Tropical Cyclogenesis in the Gulf of Mexico and Eastern Pacific: Case Study and Idealized Numerical Modeling. *J. Atmos. Sci.*, **65**, 2691–2704.
- Arora, V. K., J. F. Scinocca, G. J. Boer, J. R. Christian, K. L. Denman, G. M. Flato, V. V. Kharin, W. G. Lee, and W. J. Merryfield, 2011: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases, *Geophys. Res. Lett.*, **38**, L05805, doi:10.1029/2010GL046270.
- Ashok, K., S. K. Behera, S. A. Rao, H. Weng, and T. Yamagata, 2007: El Niño Modoki and its possible teleconnection. *J. Geophys. Res.*, **112**, C11007, doi:10.1029/2006JC003798.
- Bao, Q., and co-authors, 2012: The Flexible Global Ocean-Atmosphere-Land System model Version: FGOALS-s2. *Adv. Atmos. Sci.*, submitted.
- Barrett, B. S. and L. M. Leslie, 2009: Links between Tropical Cyclone Activity and Madden-Julian Oscillation Phase in the North Atlantic and Northeast Pacific Basins. *Mon. Weather Rev.*, **137**, 727–744. doi:10.1175/2008MWR2602.1

944 Bengtsson, L., H. Böttger, and M. Kanamitsu, 1982: Simulation of hurricane-type  
 945 vortices in a general circulation model. *Tellus*, **34**, 440-457.

946 Bengtsson, L., K. Hodges, M. Esch, N. Keenlyside, L. Kornbluh, J.-J. Luo, and T.  
 947 Yamagata, 2007: How may tropical cyclones change in a warmer climate. *Tellus*,  
 948 **59A**, 539–561.

949 Benjamini, Y., and Y. Hochberg, 1995: Controlling the false discovery rate: A practical  
 950 and powerful approach to multiple testing. *J. Roy. Stat. Soc.*, **B57**, 289-300.

951 Bi, D., M. Dix, S. Marsland, T. Hirst, S. O’Farrell and coauthors, 2012: ACCESS: The  
 952 Australian Coupled Climate Model for IPCC AR5 and CMIP5. AMOS conference,  
 953 2012, Sydney, Australia (available online at  
 954 <https://wiki.csiro.au/confluence/display/ACCESS/ACCESS+Publications>)

955 Biasutti, M., and A. H. Sobel, 2009: Delayed Sahel rainfall and global seasonal cycle in a  
 956 warmer climate. *Geophys. Res. Lett.*, **36**. doi:10.1029/2009GL041303

957 Boe, J. L., A. Hall, and X. Qi, 2009: September sea-ice cover in the Arctic Ocean  
 958 projected to vanish by 2100. *Nature Geosci.* **2**, 341–343

959 Booth, B. B. B., N. J. Dunstone, P. R. Halloran, T. Andrews, and N. Bellouin, 2012:  
 960 Aerosols implicated as a prime driver of twentieth-century North Atlantic climate  
 961 variability. *Nature*, **484**, 228-232.

962 Brekke, L.D, M. D. Dettinger, E. P. Maurer, and M. Anderson, 2008: Significance of  
 963 model credibility in estimating climate projection distributions for regional  
 964 hydroclimatological risk assessments. *Climatic Change*, **89**, 371–394 DOI  
 965 10.1007/s10584-007-9388-3

966 Burgers, G., and D. B. Stephenson, 1999: The “normality” of El Niño. *Geophys. Res.*  
967 *Lett.*, **26**, 1027–1030.

968 Camargo, S. J. and S. E. Zebiak, 2002: Improving the detection and tracking of tropical  
969 cyclones in atmospheric general circulation models. *Weather and Forecasting*, **17**,  
970 1152–1162

971 Camargo, S. J., A. G. Barnston, and S. E. Zebiak, 2005: A statistical assessment of  
972 tropical cyclone activity in atmospheric general circulation models. *Tellus*, **57A**, 589–  
973 604.

974 Camargo, S. J., A. H. Sobel, A. G. Barnston, and K. A. Emanuel, 2007: Tropical cyclone  
975 genesis potential index in climate models. *Tellus*, **59A**, 428–443.

976 Camargo, S. J., 2013: Global and regional aspects of tropical cyclone activity in the  
977 CMIP5 models. *J. Climate*, in review.

978 Cavalieri, D., C. Parkinson, P. Gloerson, and H. J. Zwally. 1997, updated 2005. /Sea ice  
979 concentrations from Nimbus-7 SMMR and DMSP SSM/I passive microwave data,  
980 June to September 2001/. Boulder, CO: National Snow and Ice Data Center. Digital  
981 media and CD-ROM.

982 Christidis, N., P. A. Stott, F. W., Zwiers, H. Shiogama, and T. Nozawa, 2010:  
983 Probabilistic estimates of recent changes in temperature: a multi-scale attribution  
984 analysis. *Clim. Dyn.*, **34**:1139–1156.

985 Colle, B. A., Z. Zhang, K. Lombardo, and P. Liu, 2012: Historical and future changes in  
986 western Atlantic extratropical cyclones in the CMIP5 models. *J. Climate*, submitted.



987 Collins, M., S. F. B. Tett, and C. Cooper, 2001: The internal climate variability of  
 988 HadCM3, a version of the Hadley Centre Coupled Model without flux adjustments.  
 989 *Climate Dynamics*, **17** (1), 61-81.

990 Compo, G. P., J. S. Whitaker, P. D. Sardeshmukh, N. Matsui, R. J. Allan, X. Yin, B. E.  
 991 Gleason, R. S. Vose, G. Rutledge, P. Bessemoulin, S. Brönnimann, M. Brunet, R. I.  
 992 Crouthamel, A. N. Grant, P. Y. Groisman, P. D. Jones, M. Kruk, A. C. Kruger, G. J.  
 993 Marshall, M. Maugeri, H. Y. Mok, Ø. Nordli, T. F. Ross, R. M. Trigo, X. L. Wang, S.  
 994 D. Woodruff, and S. J. Worley, 2011, The Twentieth Century Reanalysis Project.  
 995 *Quarterly J. Roy. Meteorol. Soc.*, **137**, 1-28. DOI: 10.1002/qj.776.

996 Cook, K. H., E. K. Vizy, Z. S. Launer, and C. M. Patricola, 2008: Springtime  
 997 intensification of the Great Plains low-level jet and Midwest precipitation in GCM  
 998 simulations of the twenty-first century. *J. Climate*, **21**, 6321–6340.

999 Corti, S., A. Weisheimer, T. Palmer, F. Doblas-Reyes, and L. Magnusson, 2012:  
 1000 Reliability of decadal predictions. *Geophys. Res. Lett.*, doi:10.1029/2012GL053354

1001 Dai, A., K. E. Trenberth, and T. R. Karl, 1999: Effects of clouds, soil moisture,  
 1002 precipitation, and water vapor on diurnal temperature range. *J. Climate*, **12**, 2451–  
 1003 2473, doi: 10.1175/1520-0442(1999)0122.0.CO;2.

1004 Delworth, T. L., and M. E. Mann, 2000: Observed and simulated multidecadal variability  
 1005 in the Northern hemisphere. *Climate Dynamics*, **16**, 661–676.

1006 de Szoeke, S. P. and C. S. Bretherton, 2005: Variability in the southerly flow into the  
 1007 eastern Pacific ITCZ. *J. Atmos. Sci.*, **62**, 4400-4411.

1008 Donner, L. J., with 28 co-authors, 2011: The dynamical core, physical parameterizations,  
 1009 and basic simulation characteristics of the atmospheric component AM3 of the GFDL  
 1010 Global Coupled Model CM3. *J. Climate*, **24**, doi:10.1175/2011JCLI3955.1.  
 1011 Dufresne, J-L., and 58 co-authors, N, 2012: Climate change projections using the IPSL-  
 1012 CM5 Earth System Model: from CMIP3 to CMIP5, *Clim. Dyn.*, submitted.  
 1013 Easterling, D.R., Briony Horton, Philip D. Jones, Thomas C. Peterson, Thomas R. Karl,  
 1014 David E. Parker, M. James Salinger, Vyacheslav Razuvayev, Neil Plummer, Paul  
 1015 Jamason and Christopher K. Folland, 1997: *Science*, **277** (5324), 364-367. DOI:  
 1016 10.1126/science.277.5324.364.  
 1017 Emanuel, K., R. Sundararajan, and J. Williams, 2008: Hurricanes and global warming:  
 1018 Results from downscaling IPCC AR4 simulations. *Bull. Amer. Meteor. Soc.*, **89**, 347-  
 1019 367.  
 1020 Enfield, D. B., A. M. Mestas-Nunez, and P. J. Trimble, 2001: The Atlantic multidecadal  
 1021 oscillation and its relationship to rainfall and river flows in the continental US.  
 1022 *Geophys. Res. Lett.*, **28**, 2077-2080.  
 1023 Fetterer, F., K. Knowles, W. Meier, and M. Savoie. 2002, updated 2009. /Sea Ice Index/.  
 1024 Boulder, Colorado USA: National Snow and Ice Data Center. Digital media.  
 1025 Folland, C. K., D. E. Parker, and T. N. Palmer, 1986: Sahel rainfall and worldwide sea  
 1026 temperatures, 1901–85. *Nature*, **320**, 602– 607.  
 1027 Furtado, J. C., E. Di Lorenzo, N. Schneider, and N. Bond, 2011: North Pacific decadal  
 1028 variability and climate change in the IPCC AR4 models. *J. Climate*, **24**, 3049-3067.  
 1029 Gent, P. R., and Coauthors, 2011: The Community Climate System Model Version 4. *J.*  
 1030 *Climate*, **24**, 4973–4991. doi: <http://dx.doi.org/10.1175/2011JCLI4083.1>

1031 Giorgi, F. and R. Francisco, 2000: Uncertainties in regional climate change prediction: a  
 1032 regional analysis of ensemble simulations with the HADCM2 coupled AOGCM.  
 1033 *Climate Dynamics*, **16** (2-3), 169–182, doi:10.1007/PL00013733.

1034 Goldenberg, S. B., C. W. Landsea, A. M. Maestas-Nunez, and W. M. Gray, 2001: The  
 1035 recent increase in Atlantic hurricane activity: Causes and implications. *Science*, **293**,  
 1036 474-479.

1037 Guilyardi, E., A. Wittenberg, A. Fedorov, M. Collins, C. Wang, A. Capotondi, G. J. van  
 1038 Oldenborgh, and T. Stockdale, 2009: Understanding El Niño in Ocean–Atmosphere  
 1039 General Circulation Models: Progress and Challenges. *Bull. Amer. Meteor. Soc.*, **90**,  
 1040 325–340.

1041 Guilyardi, E., H. Bellenger, M. Collins, S. Ferrett, W. Cai, and A. Wittenberg, 2012: A  
 1042 first look at ENSO in CMIP5. *CLIVAR Exchanges*, **17**, 29-32. ISSN: 1026-0471.

1043 Hansen, J., and Coauthors, 2005: Earth’s energy imbalance: confirmation and  
 1044 implications. *Science*, **308**, 1431-1435.

1045 Hazeleger, W., and 31 co-authors, 2010: EC-Earth: A seamless Earth system prediction  
 1046 approach in action. *Bull. Amer. Meteor. Soc.*, **91**, 1357-1363, doi:  
 1047 10.1175/2010BAMS2877.1

1048 Held, I. and M. Zhao, 2011: The response of tropical cyclone statistics to an increase in  
 1049 CO2 with fixed sea surface temperatures. *J. Climate*, **24**, 5353–5364.

1050 Held, I. M. and coauthors, 2012: High resolution AMIP simulation for CMIP5 using  
 1051 GFDL’s HIRAM model. *J. Climate*, manuscript in preparation.

1052 Higgins, R. W., A. Leetmaa, Y. Xue, and A. Barnston, 2000: Dominant factors  
 1053 influencing the seasonal predictability of U.S. precipitation and surface air  
 1054 temperature. *J. Clim.*, **13**, 3994-4017.

1055 Higgins, R. W. and W. Shi, 2001: Intercomparison of the Principal Modes of Interannual  
 1056 and Intraseasonal Variability of the North American Monsoon System. *J. Clim.*, **14**,  
 1057 403-417.

1058 Hodges, K. I., 1994: A general method for tracking analysis and its application to  
 1059 meteorological data. *Mon. Wea. Rev.*, **122**, 2573–2586.

1060 Hoerling, M., and A. Kumar, 2003: The perfect ocean for drought. *Science*, **299**, 691–  
 1061 694.

1062 Huffman, G.J., R.F. Adler, D.T. Bolvin, G. Gu, E.J. Nelkin, K.P. Bowman, Y. Hong, E.F.  
 1063 Stocker, D.B. Wolff, 2007: The TRMM Multi-satellite Precipitation Analysis:  
 1064 Quasi-Global, Multi-Year, Combined-Sensor Precipitation Estimates at Fine Scale.  
 1065 *J. Hydrometeor.*, **8** (1), 38–55.

1066 IPCC, 2007: Climate Change 2007: Synthesis Report. Contribution of working groups I,  
 1067 II, and III to the Fourth Assessment Report of the Intergovernmental Panel on  
 1068 Climate Change [Core Writing Team, Pachauri, R. K., and Reisinger, A (eds.)] IPCC,  
 1069 Geneva, Switzerland, 104 pp.

1070 Jarvinen, B. R., C. J. Neumann, and M. A. S. Davis, 1984: A tropical cyclone data tape  
 1071 for the North Atlantic Basin, 1886-1983: Contents, limitations and uses, Tech. Memo.  
 1072 NWS NHC-22, 21 pp., NOAA, Silver Spring, Md.

1073 Jiang, X. and D. E. Waliser, 2008: Northward propagation of the subseasonal variability  
1074 over the eastern Pacific warm pool. *Geophys. Res. Lett.*, **35**, L09814,  
1075 doi:10.1029/2008GL033723.

1076 Jiang, X., and D. E. Waliser, 2009: Two dominant subseasonal variability modes of the  
1077 eastern Pacific ITCZ. *Geophys. Res. Lett.*, **36**, L04704, doi: 10.1029/2008gl036820.

1078 Jiang, X., D. Waliser, D. Kim, M. Zhao, K. Sperber, W. Stern, S. Schubert, G. Zhang, W.  
1079 Wang, M. Khairoutdinov, R. Neale, and M.-I. Lee, 2011: Simulation of the  
1080 intraseasonal variability over the Eastern Pacific ITCZ in climate models. *Climate*  
1081 *Dyn.*, 10.1007/s00382-011-1098-x, 1-20.

1082 Jiang, X., E. D. Maloney, J.-L. F. Li, and D. E. Waliser, 2012: Simulations of the Eastern  
1083 North Pacific Intraseasonal Variability in CMIP5 GCMs. *J. Clim.*, 10.1175/jcli-d-12-  
1084 00526.1.

1085 Jin, F-F., S-I. An, A. Timmermann, and J. Zhao, 2003: Strong El Niño events and  
1086 nonlinear dynamical heating. *Geophys. Res. Lett.*, **30**, 1120.  
1087 doi:10.1029/2002GL016356.

1088 Jones, C. D., and others, 2011: The HadGEM2-ES implementation of CMIP5 centennial  
1089 simulations, *Geosci. Model Dev.*, **4**, 543-570, doi:10.5194/gmd-4-543-2011.

1090 Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S.  
1091 Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J.  
1092 Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetman, R. Reynolds, R. Jenne,  
1093 and D. Joseph, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Am.*  
1094 *Meteorol. Soc.*, **77**, 437–471

1095 Karl, Thomas R., George Kukla, Joyce Gavin, 1986: Relationship between Decreased  
1096 Temperature Range and Precipitation Trends in the United States and Canada, 1941–  
1097 80. *Climate Appl. Meteor.*, **25**, 1878–1886.

1098 Karl T. R., and R. W. Knight, 1998: Secular trends of precipitation amount, frequency,  
1099 and intensity in the United States. *Bull. Am. Meteorol. Soc.*, **79** (2), 231-241.

1100 Karnauskas, K. B., A. Giannini, R. Seager, and A. J. Busalacchi, 2012: A simple  
1101 mechanism for the climatological midsummer drought along the Pacific coast of  
1102 Central America. *Atmósfera*, in revision.

1103 Kavvada, A., A. Ruiz-Barradas, and S. Nigam, 2013: AMO's structure and climate  
1104 footprint in observations and IPCC AR5 climate simulations. *Climate Dynamics*,  
1105 accepted.

1106 Kayano, M. T. and V. E. Kousky, 1999: Intraseasonal (30-60 day) variability in the  
1107 global tropics: principal modes and their evolution. *Tellus Series a-Dynamic  
1108 Meteorology and Oceanography*, **51**, 373-386.

1109 Kim, D., et al., 2009: Application of MJO Simulation Diagnostics to Climate Models. *J.  
1110 Clim.*, **22**, Doi 10.1175/2009jcli3063.1, 6413-6436.

1111 Kim, D., A. H. Sobel, A. D. Del Genio, Y. Chen, S. Camargo, M.-S. Yao, M. Kelley, and  
1112 L. Nazarenko, 2012: The tropical subseasonal variability simulated in the NASA  
1113 GISS general circulation model, *J. Clim.*, in press.

1114 Kim, S. T. and J.-Y. Yu, 2012: The two types of ENSO in CMIP5 Models, *Geophysical  
1115 Research Letters*, **39**, L11704, doi:10.1029/2012GL052006.

1116 Kim, H.-M., P. J. Webster, and J. A. Curry, 2012: Evaluation of short-term climate  
 1117 change prediction in multi-model CMIP5 decadal hindcasts, *Geophys. Res. Lett.*, **39**,  
 1118 L10701, doi:10.1029/2012GL051644.

1119 Knapp, K. R., M. C. Kruk, D. H. Levinson, H. J. Diamond, and C. J. Neumann, 2010:  
 1120 The International Best Track Archive for Climate Stewardship (IBTrACS): Unifying  
 1121 tropical cyclone best track data. *Bull. Amer. Meteor. Soc.*, **91**, 363-376.  
 1122 doi:10.1175/2009BAMS2755.1

1123 Knight, J. R. et al., 2005: A signature of persistent natural thermohaline circulation cycles  
 1124 in observed climate. *Geophys. Res. Lett.*, **32**, doi:10.1029/2005GL024233.

1125 Knutson, T. R. and K. M. Weickmann, 1987: 30-60 Day Atmospheric Oscillations:  
 1126 Composite Life Cycles of Convection and Circulation Anomalies. *Mon. Weather*  
 1127 *Rev.*, **115**, 1407-1436.

1128 Knutson, T. R., J. J. Sirutis, S. T. Garner, G. A. Vecchi, and I. M. Held, 2008: Simulated  
 1129 reduction in Atlantic hurricane frequency under twenty-first-century warming  
 1130 conditions. *Nature Geoscience*, **1**, 359-364.

1131 Knutti, R., R. Furrer, C. Tebaldi, J. Cermak and G.A. Meehl, 2010, Challenges in  
 1132 combining projections from multiple models. *J. Climate*, **23**, 2739-2758,  
 1133 doi:10.1175/2009JCLI3361.1

1134 Knutti, R., 2010: The end of model democracy? *Climatic Change*, **102**, 395–404. DOI  
 1135 10.1007/s10584-010-9800-2.

1136 Kug, J.-S., F.-F. Jin, and S.-I. An, 2009: Two-types of El Niño events: Cold Tongue El  
 1137 Niño and Warm Pool El Niño. *J. Climate*, **22**, 1499-1515.

1138 Kumar, S., V. Merwade, J. Kam, and K. Thurner, 2009: Streamflow trends in Indiana:  
 1139 effects of long term persistence, precipitation and subsurface drains, *J. Hydrology*,  
 1140 **374** (1-2), 171–183.

1141 Kumar, S., J. Kinter III, P. Dirmeyer, Z. Pan, and J. Adams, 2012: Multi-decadal Climate  
 1142 Variability and the "Warming Hole" in North America - results from CMIP5 20th and  
 1143 21st Century Climate Simulations. *J. Climate*. doi:10.1175/JCLI-D-12- 00535.1, in  
 1144 press.

1145 Kumar, S., V. Merwade, J. Kinter III, and D. Niyogi, 2013: Evaluation of Temperature  
 1146 and Precipitation Trends and Long-term Persistence in CMIP5 20<sup>th</sup> Century Climate  
 1147 Simulations. *J. Climate*. doi:10.1175/JCLI-D-12-00259.1, in press.

1148 Larkin, N. K. and D. E. Harrison, 2005: On the definition of El Niño and associated  
 1149 seasonal average U.S. weather anomalies. *Geophys. Res. Lett.*, **32**:L13705.  
 1150 doi:10.1029/2005GL022738.

1151 Lauritsen, R., and J. Rogers, 2012: U.S. diurnal temperature range variability and  
 1152 regional causal mechanisms, 1901-2002. *J. Climate*, doi:10.1175/JCLI-D-11-00429.1,  
 1153 in press.

1154 Lee, T., and M. J. McPhaden, 2010: Increasing intensity of El Niño in the central-  
 1155 equatorial Pacific. *Geophys. Res. Lett.*, **37**, L14603, doi:10.1029/2010GL044007.

1156 Lettenmaier D. P., E. F. Wood, J. R. Wallis, 1994: Hydro-climatological trends in the  
 1157 continental United States, 1948-88. *J. Climate*, **7**, 586-607.

1158 Liang, J., X.-Q. Yang, and D.-Z. Sun, 2012: The effect of ENSO events on the Tropical  
 1159 Pacific Mean Climate: Insights from an Analytical Model. *J. Climate*, Accepted.



1160 Livezey, R. E., and W. Y. Chen, 1983: Statistical field significance and its determination  
 1161 by Monte Carlo techniques. *Mon. Wea. Rev.*, **111**, 46–59.  
 1162 Lorenz, D. J. and D. L. Hartmann, 2006: The Effect of the MJO on the North American  
 1163 Monsoon. *J. Clim.*, **19**, 333-343.  
 1164 Magaña, V., J.A. Amador, and S. Medina, 1999: The Midsummer Drought over Mexico  
 1165 and Central America. *J. Climate*, **12**, 1577–1588.  
 1166 Maloney, E. D. and D. L. Hartmann, 2000a: Modulation of Eastern North Pacific  
 1167 Hurricanes by the Madden-Julian Oscillation. *J. Clim.*, **13**, 1451-1460.  
 1168 Maloney, E. D. and D. L. Hartmann, 2000b: Modulation of hurricane activity in the Gulf  
 1169 of Mexico by the Madden-Julian oscillation. *Science*, **287**, 2002-2004.  
 1170 Maloney, E. D. and S. K. Esbensen, 2003: The Amplification of East Pacific Madden-  
 1171 Julian Oscillation Convection and Wind Anomalies during June-November. *J. Clim.*,  
 1172 **16**, 3482-3497.  
 1173 Maloney, E. D. and S. K. Esbensen, 2005: A Modeling Study of Summertime East  
 1174 Pacific Wind-Induced Ocean-Atmosphere Exchange in the Intraseasonal Oscillation.  
 1175 *J. Clim.*, **18**, 568-584.  
 1176 Maloney, E. D. and S. K. Esbensen, 2007: Satellite and Buoy Observations of Boreal  
 1177 Summer Intraseasonal Variability in the Tropical Northeast Pacific. *Mon. Weather*  
 1178 *Rev.*, **135**, 3-19.  
 1179 Maloney, E. D., D. B. Chelton, and S. K. Esbensen, 2008: Subseasonal SST variability in  
 1180 the tropical eastern north Pacific during boreal summer. *J. Clim.*, **21**, Doi  
 1181 10.1175/2007jcli1856.1, 4149-4167.

1182 Maloney, E. D., S. J. Camargo, E. Chang, B. Colle, R. Fu, K. L. Geilw, Q. Hu, X. Jiang,  
 1183 N. Johnson, K. B. Karneauskas, J. Kinter, B. Kirtman, S. Kumar, B. Langenbrunner,  
 1184 K. Lombardo, L. Long, A. Mariotti, J. E. Meyerson, K. Mo, J. D. Neelin, Z. Pan, R.  
 1185 Seager, Y. Serraw, A. Seth, J. Sheffield, J. Thibeault, S.-P. Xie, C. Wang, B. Wyman,  
 1186 and M. Zhao, 2011: North American Climate in CMIP5 Experiments: Part III:  
 1187 Assessment of 21<sup>st</sup> Century Projections. *J. Climate*, submitted.  
 1188 Manabe, S., J. L. Holloway, and H. M. Stone, 1970: Tropical circulation in a time-  
 1189 integration of a global model atmosphere. *J. Atmos. Sci.*, **27**, 580-613.  
 1190 Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. Francis, 1997: A Pacific  
 1191 interdecadal climate oscillation with impacts on salmon production. *Bull. Amer.*  
 1192 *Meteor. Soc.*, **78**, 1069-1079.  
 1193 Martin, E. R. and C. Schumacher, 2010: Modulation of Caribbean Precipitation by the  
 1194 Madden-Julian Oscillation. *J. Clim.*, doi:10.1175/2010JCLI3773.1, in press.  
 1195 McCabe, G. J., and D. M. Wolock, 2002: A step increase in streamflow in the  
 1196 conterminous United States. *Geophys. Res. Lett.*, **29** (24), 2185, doi:10.1029/  
 1197 2002GL015999.  
 1198 McCabe, G. J., M. A. Palecki, and J. L. Betancourt, 2004: Pacific and Atlantic Ocean  
 1199 influences on multidecadal drought frequency in the United States. *Proc Natl Acad*  
 1200 *Sci.*, **101**, 4136-4141.  
 1201 Magana, V., J. A. Amador, and S. Medina, 1999: The midsummer drought over Mexico  
 1202 and Central America. *J. Clim.*, **12**, 1577-1588.

1203 McCabe, G. J., J. L. Betancourt, S. T. Gray, M. A. Palecki, and H. G. Hidalgo, 2008:  
 1204 Associations of multi-decadal sea-surface temperature variability with US drought.  
 1205 *Quat. Int.*, **188**, 31–40.  
 1206 Medhaug, I., and T. Furevik, 2011: North Atlantic 20th century multidecadal variability  
 1207 in coupled climate models: sea surface temperature and ocean overturning circulation.  
 1208 *Ocean Sci.*, **7**, 389-404.  
 1209 Merryfield, W. J., W.-S. Lee, G. J. Boer, V. V. Kharin, J. F. Scinocca, G. M. Flato, and S.  
 1210 Polavarapu, 2012: The Canadian seasonal to interannual prediction system. Part I:  
 1211 Models and initialization. *Monthly Weather Review*, submitted.  
 1212 Molinari, J. and D. Vollaro, 2000: Planetary- and synoptic-scale influences on eastern  
 1213 Pacific tropical cyclogenesis. *Mon. Weather Rev.*, **128**, 3296-3307.  
 1214 Muller, W. A., and E. Roeckner, 2008: ENSO teleconnections in projections of future  
 1215 climate in ECHAM5/MPI-OM. *Clim. Dyn.*, **31**, 533–549.  
 1216 Nigam, S., B. Guan, and A. Ruiz-Barradas, 2011: Key role of the Atlantic Multidecadal  
 1217 Oscillation in 20th century drought and wet periods over the Great Plains. *Geophys.*  
 1218 *Res. Lett.*, **38**, L16713, doi:10.1029/2011GL048650.  
 1219 Ogata, T., S.-P. Xie, A. Wittenberg, and D.-Z. Sun, 2012: Interdecadal amplitude  
 1220 modulation of El Nino/Southern Oscillation and its impacts on tropical Pacific  
 1221 decadal variability. *J. Climate*, submitted.  
 1222 Oouchi, K., J. Yoshimura, H. Yoshimura, R. Mizuta, S. Kusunoki, and A. Noda, 2006:  
 1223 Tropical cyclone climatology in a global-warming climate as simulated in a 20 km  
 1224 mesh global atmospheric model: Frequency and wind intensity analysis. *J. Meteor.*  
 1225 *Soc. Japan*, **84**, 259–276.

1226 Ottera, O. H., and Coauthors, 2003: The sensitivity of the present-day Atlantic meridional  
 1227 overturning circulation to freshwater forcing. *Geophys. Res. Lett.*, **30**, 1898,  
 1228 doi:10.1029/2003GL017578.

1229 Osgood, D., A. Giannini, J. Hansen, E. Holthaus, A. Ines, Y. Kaheil, K. Karanaskas, M.  
 1230 McLaurin, A. Robertson, K. Shirley, and M. Vicarelli, 2009: Designing index-based  
 1231 weather insurance for farmers in Central America, Final report to the Commodity  
 1232 Risk Management Group, ARD, World Bank. Tech. Rep. No. 09-01. International  
 1233 Research Institute for Climate and Society, Palisades, NY, USA, 79 pp.

1234 Oshima, K., and Y. Tanimoto, 2009: An evaluation of reproducibility of the Pacific  
 1235 Decadal Oscillation in the CMIP3 simulations. *J. Meteor. Soc. Japan*, **87**, 755-770.

1236 Pan, Z., X. Liu, S. Kumar, Z. Gao, and J. Kinter, 2012: Inter-model variability and  
 1237 mechanism attribution of central and southeastern U.S. anomalous cooling in the 20<sup>th</sup>  
 1238 century as simulated by CMIP5 models. *J. Clim.* (under review)

1239 Portig, W. H., 1961: Some climatological data of Salvador, Central America. *Weather*,  
 1240 **16**, 103–112.

1241 Portmann, R. W., S. Solomon, and G. C. Hegel, 2009: Spatial and seasonal patterns in  
 1242 climate change, temperatures, and precipitation across the United States. *Proc. Natl.*  
 1243 *Acad. Sci.*, 106, 7324–7329.

1244 Räisänen J, 2007 How reliable are climate models?. *Tellus, Ser. A, Dynam. Meteorol.*  
 1245 *Oceanogr.* 59, 2–29. doi:10.1111/j.1600-0870.2006.00211.x.

1246 Rayner, R., D. Parker, E. Horton, C. Folland, L. Alexander, and D. Rowel, 2003: Global  
 1247 analyses of sea surface temperature, sea ice, and night marine air temperature since

1248 the late nineteenth century. *J. Geophys. Res.*, **108**, D14,4407,doi:10.1029/2002JD002  
1249 670.

1250 Rauscher, S. A., F. Giorgi, N. S. Diffenbaugh, and A. Seth, 2008: Extension and  
1251 intensification of the Meso–American mid–summer drought in the 21st century,  
1252 *Climate Dynamics*, **31**(5), 551–571.

1253 Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K.  
1254 Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D.  
1255 Lohmann, and D. Toll, 2004: The Global Land Data Assimilation System. *Bull.*  
1256 *Amer. Met. Soc.*, 381–394.

1257 Rodgers, K. B., P. Friederichs, and M. Latif, 2004: Tropical Pacific decadal variability  
1258 and its relation to decadal modulation of ENSO. *J. Climate*, **17**, 3761–3774.

1259 Rowell, D. P., C. K. Folland, K. Maskell, and M. N. Ward, 1995: Variability of summer  
1260 rainfall over tropical North-Africa (1906–92) observations and modelling. *Q J R*  
1261 *Meteorol. Soc.*, **121**, 669–704.

1262 Ruiz-Barradas, A., and S. Nigam, 2005: Warm season precipitation variability over the  
1263 US Great Plains in observations, NCEP and ERA-40 reanalyses, and NCAR and  
1264 NASA atmospheric model simulations. *J. Climate*, **18**, 1808–1830.

1265 Ruiz-Barradas, A., and S. Nigam, 2006: IPCC’s twentieth-century climate simulations:  
1266 Varied representations of North American hydroclimate variability. *J. Climate*, **19**,  
1267 4041–4058

1268 Ruiz-Barradas, A., and S. Nigam, 2010: SST–North American hydroclimate links in  
1269 AMIP simulations of the Drought Working Group models: A proxy for the idealized  
1270 drought modeling experiments. *J. Climate*, **23**, 2585–2598.

1271 Ruiz-Barradas, A., S. Nigam, and A. Kavvada, 2013: The Atlantic Multidecadal  
 1272 Oscillation in 20<sup>th</sup> century climate simulations: uneven progress from CMIP3 to  
 1273 CMIP5. Under review in *Climate Dynamics*.  
 1274 Rydbeck, A. V., E. D. Maloney, S.-P. Xie, J. Hafner, and J. Shaman, 2012: Remote  
 1275 Forcing Versus Local Feedback of East Pacific Intraseasonal Variability during  
 1276 Boreal Summer. *J. Clim.*, 10.1175/jcli-d-12-00499.1.  
 1277 Sakamoto, T. T., Y. Komuro, T. Nishimura, M. Ishii, H. Tatebe, H. Shiogama, A.  
 1278 Hasegawa, T. Toyoda, M. Mori, T. Suzuki, Y. Imada, T. Nozawa, K. Takata, T.  
 1279 Mochizuki, K. Ogochi, S. Emori, H. Hasumi and M. Kimoto, 2012: MIROC4h - a  
 1280 new high-resolution atmosphere-ocean coupled general circulation model. *J. Meteor.*  
 1281 *Soc. Japan*, **90** (3), 325-359.  
 1282 Santer, B. D. *et al.*, 2000: Statistical significance of trends and trend differences in layer-  
 1283 average atmospheric temperature time series. *J. Geophys. Res.* **105**, 7337-7356.  
 1284 Schopf, P. S., and R. J. Burgman, 2006: A simple mechanism for ENSO residuals and  
 1285 asymmetry. *J. Climate*, **19**, 3167–3179.  
 1286 Schubert, S. D., M. J. Suarez, P. J. Pegion, R. D. Koster, and J. T. Bacmeister, 2004:  
 1287 Causes of long-term drought in the U.S. Great Plains. *J. Climate*, **17**, 485–503.  
 1288 Schubert, S., et al., 2009: A USCLIVAR project to assess and compare the responses of  
 1289 global climate models to drought-related SST forcing patterns: Overview and results.  
 1290 *J. Climate*, **22**, 5251–5272, doi:10.1175/2009JCLI3060.1.  
 1291 Serra, Y. L., G. N. Kiladis, and K. I. Hodges, 2010: Tracking and Mean Structure of  
 1292 Easterly Waves Over the Intra-Americas Sea. *J. Clim.*, **0**,  
 1293 doi:10.1175/2010JCLI3223.1.

1294 Seth, A., S. Rauscher, M. Rojas, S. Camargo, A. Giannini, 2011: Enhanced spring  
1295 convective barrier for monsoons in a warmer world? *Climatic Change Let.*, **104**,  
1296 403~H414, doi:10.1007/s10584-010- 9973-8.

1297 Seth, A., M. Rojas, and S. A. Rauscher, 2010: CMIP3 projected changes in the annual  
1298 cycle of the South American monsoon. *Clim. Change*, **98**: 331~H357.  
1299 doi:10.1007/s10584-009-9736-6

1300 Sheffield, J., A. Barrett, B. Colle, R. Fu, K. L. Geil, Q. Hu, J. Kinter, S. Kumar, B.  
1301 Langenbrunner, K. Lombardo, L. N. Long, E. Maloney, A. Mariotti, J. E. Meyerson,  
1302 K. C. Mo, J. D. Neelin, Z. Pan, A. Ruiz-Barradas, Y. L. Serra, A. Seth, J. M.  
1303 Thibeault, and J. C. Stroeve, 2012: North American Climate in CMIP5 Experiments.  
1304 Part I: Evaluation of 20th Century Continental and Regional Climatology. *J. Climate*,  
1305 submitted.

1306 Small, R. J., S. P. De Szoeko, and S.-P. Xie, 2007: The Central American midsummer  
1307 drought: Regional aspects and large-scale forcing. *J. Clim.*, **20**, Doi  
1308 10.1175/Jcli4261.1, 4853-4873.

1309 Smith, T. M., R.W. Reynolds, T. C. Peterson, and J. Lawrimore, 2008: Improvements to  
1310 NOAA's historical merged land-ocean surface temperature analysis (1880-2006). *J*  
1311 *Clim.*, **21**, 2283-2296.

1312 Smith, T. M., R. W. Reynolds, T. C. Peterson, and J. Lawrimore, 2008: Improvements to  
1313 NOAA's Historical Merged Land-Ocean Surface Temperature Analysis (1880-2006).  
1314 *J. Climate*, **21**, 2283-2296.

1315 Sun, D.-Z., 2000a: Global climate change and ENSO: a theoretical framework. El Niño:  
1316 Historical and Paleoclimatic Aspects of the Southern Oscillation, Multiscale

1317 variability and Global and Regional Impacts .443-463. Cambridge: Cambridge  
 1318 University Press, edited by Diaz H. F. and V. Markgraf, 476 pp.

1319 Sun, D.-Z., and T. Zhang, 2006: A regulatory effect of ENSO on the time-mean thermal  
 1320 stratification of the equatorial upper ocean. *Geophys. Res. Lett.*, **33**, L07710.  
 1321 doi:10.1029/2005GL025296.

1322 Sun, D.-Z., and F.O. Bryan, 2010: Preface. "Climate Dynamics: Why Does Climate  
 1323 Vary?". AGU Geophysical Monograph, Edited by D.-Z. Sun and F. Bryan, AGU.

1324 Sun, D.-Z., 2010: The Diabatic and Nonlinear Aspects of El Nino Southern Oscillation:  
 1325 Implications for its Past and Future Behavior. page 79-104. "Climate Dynamics: Why  
 1326 Does Climate Vary?". AGU Geophysical Monograph, Edited by D.-Z. Sun and F.  
 1327 Bryan, AGU.

1328 Sun, Y., D.-Z. Sun, and L. Wu, 2012: Evidence from climate models for a role of ENSO  
 1329 events in shaping the climatological size and temperature of the warm-pool.  
 1330 *Advances in Atmospheric Sciences*, Accepted.

1331 Sun, D.-Z., T. Zhang, Y. Sun, and Y. Yu, 2012: Rectification of El Nino-Southern  
 1332 Oscillation into Climate Anomalies of Decadal and Longer Time-scales. *J. Climate*,  
 1333 Submitted

1334 Sun, F., and J-Y. Yu, 2009: A 10–15-year modulation cycle of ENSO intensity. *J.*  
 1335 *Climate*, **22**, 1718–1735.

1336 Sutton R., and D. Hodson, 2005: Atlantic Ocean Forcing of North American and  
 1337 European Summer Climate. *Science*, **309**, 115-117 doi: 10.1126/science.1109496.



1338 Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the  
 1339 experiment design, *Bull. Am. Meteorol. Soc.*, **93**, 485–498, doi:10.1175/BAMS-D-11-  
 1340 00094.1.

1341 Tebaldi, C., J. M. Arblaster, and R. Knutti, 2011: Mapping model agreement on future  
 1342 climate projections. *Geophys. Res. Lett.*, **38**, L23701, doi: 10.1029/2011GL049863.

1343 Ting, M., Y. Kushnir, R. Seager, and C. Li, 2009: Forced and internal twentieth-century  
 1344 SST trends in the North Atlantic. *J. Climate*, **22**, 1469–1481.

1345 Trenberth, K. E., D. P. Stepaniak, 2001: Indices of El Niño evolution. *J. Climate*, **14**,  
 1346 1697-1701.

1347 Universidad Nacional Autónoma de México (UNAM), 2007: Gridded precipitation and  
 1348 temperature analysis from the Centro de Ciencias de la Atmósfera, Mexico; available  
 1349 from the International Research Institute for Climate and Society  
 1350 (<http://ingrid.ldeo.columbia.edu/SOURCES/UNAM/gridded/monthly/v0705/>).

1351 van Oldenborgh, G. J., S. Philip, and M. Collins, 2005: El Niño in a changing climate: A  
 1352 multi-model study. *Ocean Sci.*, **1**, 81-95.

1353 Vecchi, G. A., and B. J. Soden, 2007: Increased tropical Atlantic wind shear in model  
 1354 projections of global warming. *Geophys. Res. Lett.*, **34**, 10.1029/2006GL028905.

1355 Vecchi, G. A., et al., 2011: Statistical-dynamical predictions of seasonal North Atlantic  
 1356 hurricane activity. *Mon. Wea. Rev.*, **139**, 1070–1082.

1357 Veres, M. C., and Q. Hu, 2013: AMO-forced regional processes affecting summertime  
 1358 precipitation variations in the central United States. *J. Climate*, in press.

1359 Vintzileos, A., M. M. Rienecker, M. J. Suarez, S. D. Schubert, and S. K. Miller, 2005:  
1360 Local versus remote wind forcing of the equatorial Pacific surface temperature in July  
1361 2003. *Geophys. Res. Lett.*, **32**, L05702, doi:10.1029/2004gl02197.

1362 Voltaire, A., and others, 2012: The CNRM-CM5.1 global climate model: Description  
1363 and basic evaluation, *Clim. Dyn.*, doi:10.1007/s00382-011-1259-y, in press.

1364 Volodin, E. M., N. A. Diansky, and A. V. Gusev, 2010: Simulating Present-Day Climate  
1365 with the INMCM4.0 Coupled Model of the Atmospheric and Oceanic General  
1366 Circulations. *Izvestia, Atmospheric and Oceanic Physics*, **46**, 414-431

1367 Vose, R. S., D. R. Easterling, and B. Gleason, 2005: Maximum and minimum  
1368 temperature trends for the globe: An update through 2004, *Geophys. Res. Lett.*, **32**,  
1369 L23822, doi:10.1029/2005GL024379.

1370 Walsh, K., S. Lavender, H. Murakami, E. Scoccimarro, L.-P. Caron and M. Ghantous,  
1371 2010: The Tropical Cyclone Climate Model Intercomparison Project. *Hurricanes and*  
1372 *Climate Change*, **2**, Springer, 1-24.

1373 Wang, C., and D. B. Enfield, 2001: The tropical Western Hemisphere warm pool.  
1374 *Geophys. Res. Lett.*, **28**, 1635–1638.

1375 Wang, C., D. B. Enfield, S.-K. Lee, and C. W. Landsea, 2006: Influences of the Atlantic  
1376 warm pool on Western Hemisphere summer rainfall and Atlantic hurricanes. *J. Clim.*,  
1377 **19**, 3011-3028.

1378 Wang, C. and R. H. Weisberg, 2000: The 1997-98 El Niño evolution relative to previous  
1379 El Niño events. *J. Climate*, **13**, 488-501.

1380 Wang, C., and S.-K. Lee, 2009: Co-variability of Tropical Cyclones in the North Atlantic  
 1381 and the Eastern North Pacific. *Geophys. Res. Lett.*, **36**, L24702,  
 1382 doi:10.1029/2009GL041469.

1383 Wang, C., S. Dong, A. T. Evan, G.R. Foltz, and S.-K. Lee, 2012: Multidecadal  
 1384 covariability of North Atlantic sea surface temperature, African dust, Sahel rainfall  
 1385 and Atlantic hurricanes. *J. Climate*, **25**, 5404-5415.

1386 Watanabe, M., and Coauthors, 2010: Improved Climate Simulation by MIROC5: Mean  
 1387 States, Variability, and Climate Sensitivity. *J. Climate*, **23**, 6312–6335.

1388 Weaver, S., A. Ruiz-Barradas, and S. Nigam, 2009: Pentad evolution of the 1988 drought  
 1389 and 1993 flood over the Great Plains: A NARR Perspective on the atmospheric and  
 1390 terrestrial water balance. *J. Climate*, **22**, 5366-5384.

1391 Weaver, S. J., and S. Nigam, 2008: Variability of the Great Plains low-level jet: Large-  
 1392 scale circulation context and hydroclimate impacts. *J. Climate*, **20**, 1532–1551.

1393 Wilks, D. S., 2006: On “field significance” and the false discovery rate. *J. Appl.*  
 1394 *Meteorol. Climatol.*, **45**, 1181-1189.

1395 Wu, Q., 2010: Associations of diurnal temperature range change with the leading climate  
 1396 variability modes during the Northern Hemisphere wintertime and their implication  
 1397 on the detection of regional climate trends *J. Geophys. Res.*, **115**, D19101,  
 1398 doi:10.1029/2010JD014026

1399 Wu, M. L. C., S. D. Schubert, M. J. Suarez, and N. E. Huang, 2009: An Analysis of  
 1400 Moisture Fluxes into the Gulf of California. *J. Clim.*, **22**, Doi 10.1175/2008jccli2525.1,  
 1401 2216-2239.

1402 Xie, P., and P.A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on  
 1403 gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer.*  
 1404 *Meteor. Soc.*, **78**, 2539 - 2558.

1405 Xin X., Wu T., Zhang J., 2012: Introductions to the CMIP 5 simulations conducted by the  
 1406 BCC climate system model (in Chinese). *Advances in Climate Change Research*.  
 1407 submitted.

1408 Yu, J.-Y. and H.-Y. Kao, 2007: Decadal Changes of ENSO Persistence Barrier in SST  
 1409 and Ocean Heat Content Indices: 1958-2001. *J. Geophys. Res.*, **112**, D13106,  
 1410 doi:10.1029/2006JD007654.

1411 Yu., J.-Y. and S. T. Kim, 2010: Identification of Central-Pacific and Eastern-Pacific  
 1412 Types of ENSO in CMIP3 Models, *Geophys. Res. Lett.*, **37**, L15705,  
 1413 doi:10.1029/2010GL044082.

1414 Yu., J.-Y., Y. Zou, S. T. Kim, and T. Lee, 2012: The Changing Impact of El Niño on US  
 1415 Winter Temperatures. *Geophys. Res. Lett.*, doi:10.1029/2012GL052483.

1416 Yu, J.-Y. and Y. Zou, 2013: The enhanced drying effect of Central-Pacific El Nino on US  
 1417 winter, *Environmental Research Letters*, 8, doi:10.1088/1748-9326/8/1/014019.

1418 Yukimoto, S., et al., 2012: A new global climate model of the Meteorological Research  
 1419 Institute: MRI-CGCM3—Model description and basic performance, *J. Meteorol. Soc.*  
 1420 *Jpn.*, **90a**, 23–64.

1421 Zanchettin, D., A. Rubino, D. Matei, O. Bothe, and J. H. Jungclaus, 2012: Multidecadal-  
 1422 to-centennial SST variability in the MPI-ESM simulation ensemble for the last  
 1423 millennium. *Clim. Dyn.*, **39**, 419-444 doi:10.1007/s00382-012-1361-9.

1424 Zhang, T., D.-Z. Sun, R. Neale, and P. J. Rasch, 2009: An Evaluation of ENSO  
 1425 Asymmetry in the Community Climate System Models: A View from the Subsurface.  
 1426 *J. Climate*, **22**, 5933-5961.

1427 Zhang, T., and D.-Z. Sun, 2012: An Evaluation of ENSO Asymmetry in CMIP5 Models.  
 1428 *J. Climate* (in preparation)

1429 Zhang X., L. A. Vincent, W. D. Hogg, and A. Nittsoo (2000), Temperature and  
 1430 precipitation trends in Canada during 20th century, *Atmosphere-Ocean*, **38** (3), 395-  
 1431 429.

1432 Zhang, Y., J. M. Wallace, and D. S. Battisti, 1997: ENSO-like interdecadal variability:  
 1433 1900-93. *J. Climate*, **10**, 1004-1020.

1434 Zhang, Z. S., Nisancioglu, K., Bentsen, M., Tjiputra, J., Bethke, I., Yan, Q.,  
 1435 Risebrobakken, B., Andersson, C., and Jansen, E., 2012: Pre-industrial and mid-  
 1436 Pliocene simulations with NorESM-L. *Geosci. Model Dev.*, **5**, 523-533,  
 1437 doi:10.5194/gmd-5-523-2012.

1438 Zhao, M., I. M. Held, S.-J. Lin, and G. A. Vecchi, 2009: Simulations of global hurricane  
 1439 climatology, interannual variability, and response to global warming using a 50km  
 1440 resolution GCM. *J. Climate*, **22**, 6653–6678.

1441 Zhao, M., I. M. Held, and G. A. Vecchi, 2010: Retrospective forecasts of the hurricane  
 1442 season using a global atmospheric model assuming persistence of SST anomalies.  
 1443 *Mon. Wea. Rev.*, **138**, 3858–3868.

1444 Zhou, L. R. E. Dickinson, A. Dai and P. Dirmeyer, 2008: Detection and attribution of  
 1445 anthropogenic forcing to diurnal temperature range changes from 1950 to 1999:

1446 comparing multi-model simulations with observations. *Clim. Dynam.*, **35**, 7-8, 1289-  
1447 1307, DOI: 10.1007/s00382-009-0644-2.  
1448

## Figure Captions

**Figure 1.** Taylor diagram for summer mean (May-September) rainfall over the eastern Pacific (150°W-80°W; 5°S-30°N) simulated in CMIP5 GCMs. The rainfall observations are based on TMPA data.

**Figure 2.** Spatial distribution of amplitude (a) and phase (b) of the first leading complex EOF (CEOF1) mode based on 30-90-day band-pass filtered TRMM rainfall during boreal summer (June-September) over the eastern Pacific. To make the spatial phase patterns of the CEOF1 based on the observations and simulations comparable to each other, the spatial phase of CEOF1 for each dataset is adjusted by setting the domain averaged value to be zero over a small box region of 110°W-100°W, 10-15°N. Contours are only displayed where the local variance explained by CEOF1 exceeds 8%; (c): X-axis: Pattern correlation coefficients of the CEOF1 mode between TRMM observations and CMIP5 GCM simulations. Y-axis: Relative amplitudes of CEOF1 in model simulations to their observed counterparts. Both pattern correlations and amplitudes are derived by averaging over the area of 5°N-25°N, 140°W-80°W where the active ISV is observed. The black “star” mark represents the TMPA observations.

**Figure 3.** Summertime (June-September) MSD strength (mm/day) for three observational estimates (TRMM 3B43, UNAM and GPCP) and the CMIP5 MME mean for 23 models (see Table 3). Also shown (bottom row) are the MME standard deviation and histogram of the pattern correlations between individual models and the MME mean. All model output and observational data were regridded onto a common 0.5° grid.

**Figure 4.** Storm track density (top) and mean strength (bottom) for ERA Interim and seven CMIP5 models (CanESM2, CCSM4, GFDL-ESM2M, HadGEM2-ES, MIROC5, MPI-ESM-L and MRI-CGCM3). Tracks are based on 6-hourly 850hPa relative vorticity smoothed to T42 spatial resolution to better capture the synoptic features of the vorticity field.

**Figure 5.** Tracks of tropical cyclone-like storms in the CMIP5 historical runs in the period 1950-2005 (GFDL-ESM2M (1 ensemble member), HadGEM2 (1), MPI-ESM-LR (3), MRI CGCM3 (5) and MIROC5 (1) models) and in observations for the same period. The number of storms in each case is given in the bottom right corner of each panel. One ensemble member is used for each model.

**Figure 6.** Mean number of TCs per month in models (GFDL-ESM2M, HadGEM2-ES (in the figure HGEM2), MPI-ESM-LR, MRI-CGCM3, MIROC5) and observations in the North Atlantic (top left panel) and eastern North Pacific (top right panel), using only ensemble 1 for MRI-CGCM3. Number of TCs per year in the period 1950-2005 in models and observations for the North Atlantic (bottom left panel) and eastern North Pacific (bottom right panel). The blue box shows the 25-75 percentile range, with the median shown as a red line. The whiskers and red crosses show the data outside of middle quartiles.

**Figure 7.** Upper panels: Comparison of observed and C180HIRAM (one realization)

simulated hurricane tracks for the N. Atlantic and E. Pacific for 1981-2008. Middle panel: Comparison of observed and C180HIRAM simulated annual hurricane count statistics. Blue boxes show the 25-75 percentile range, with the median shown as a red line and the mean shown as a red star. The whiskers show the maximum and minimum values. The annual statistics are computed based on a 3-member ensemble mean for 1981-2008. Lower panels: Observed and model simulated seasonal cycle (number of hurricanes per month) for the N. Atlantic and E. Pacific from the 3-member ensemble mean (1=JAN, 12=DEC).

**Figure 8. Taylor diagrams for (a) El Niño and (b) La Niña composite 300 hPa geopotential patterns over the region from East Asia – North America.** Higher performance (pattern correlation > 0.6, RMS difference < 13m in both (a) and (b)) models are indicated in red, whereas lower performance models are indicated in blue. In (a) HadCM3, which falls outside of the plot, has a pattern correlation of -0.3 and RMS difference of 17.6 m. The points labeled “ens” in red, blue, and green represent the higher performance, lower performance, and total ensemble, respectively. The composites are normalized by the Niño 3.4 SST amplitude to focus on pattern differences independent of ENSO amplitude differences. The observational reference is based on the NCEP/NCAR reanalysis for 1950-2010, whereas the CMIP calculations are based on the full historical period (1850-2005) for one run of each model.

**Figure 9.** Composites of (a,b,c) 300 hPa height (z300, m), (d,e,f) near surface air temperature (SAT, °C), (g,h,i) precipitation (Precip, mm/day), and (j,k,l) SST (°C) anomalies during DJF El Niño episodes in observations (left column), and both high (middle column) and low performance CMIP5 ensembles (right column) described in Fig. 8. The observational SAT and precipitation composites are based on the CRU TS3.1 land near-surface temperature and precipitation datasets for 1901-2009. The z300, SAT, and Precip composites are normalized by the Niño 3.4 SST anomaly. Stippling in the observed (a) z300, (d) SAT, and (g) precipitation composites indicates anomalies that are statistically significant at the 5% level.

**Figure 10.** Composite DJF El Niño (a) precipitation (mm/day) and (b) SST (°C) difference between the high and low performance CMIP5 ensemble described in Fig. 8. Stippling indicates differences that are statistically significant at the 5% level. (c) DJF SST climatology difference (°C) between the high and low performance ensemble, and (d) high and (e) low performance SST climatology bias (°C) for the 1951-2000 period.

**Figure 11.** US winter surface air temperature regressed on the EP (top six rows) and CP (bottom six rows) ENSO indices from observations and the CMIP5 models. Observational air temperature data are from the NCEP-NCAR reanalysis and SSTs are from the ERSST dataset for 1950-2010.

**Figure 12.** (a) Scatter plot of pattern correlations between the regression patterns from the CMIP5 models and those from the observations (NCEP-NCAR reanalysis and HadISST dataset) for EP versus CP ENSO; (b) Scatter plot of the intensities of the EP/CP ENSO from the CMIP5 models and the observation (ERSST). The values shown are the maximum standard deviations of the EOF patterns of the two types of the ENSO



calculated using a regression-EOF method. The blue dashed lines indicate the lower limit of the 95% significance interval of the observed ENSO intensities based on an F-test.

**Figure 13.** The sum of the composite SST anomalies between the two phases of ENSO from the HadISST observations and CMIP5 coupled models. The definition of the warm phase and cold phase of ENSO follows that of Zhang et al. (2009). The length of data used in the calculation is 50 years for all the models and observations (1950-99).

**Figure 14.** The standard deviation (upper panel) and skewness (lower panel) of monthly Niño-3 SST anomalies from observations and CMIP5 model simulations. The length of data for computing the standard deviation and skewness is 50 years for the observations (1950–99). For the model, the standard deviation and skewness were calculated for a 50-year moving window over 100 years of the model run for a total of 601 samples. The figure shows the mean of the samples and +- the standard deviation across the samples. Data used are the same as for Figure 13.

**Figure 15.** Summertime wet and dry circulation patterns for the central US from the NCEP/NCAR reanalysis. a) and b) show, respectively, summertime precipitation anomalies (contours) in wet and dry years, in reference to the Great Plains precipitation, and the vertically integrated moisture fluxes from the surface to the top of the troposphere (arrows). c) The differences between a) and b). d) and e) show the corresponding 850hPa geopotential height (countour) and 925hPa wind anomalies (arrows) for the wet and dry summer, respectively. Their differences are summerized in f).

**Figure 16.** Same as Fig. 15 but for CCSM4 simulation results.

**Figure 17.** PDO SST patterns in observations and CMIP5 models. Linear regression of SST on the PDO index in (a) observations and (b) the CMIP5 ensemble, and (c) the CMIP5 minus observed PDO regression. Observations are from the HadISST dataset for the period between 1870 and 2009. For the CMIP5 models, the analysis period begins as early as 1850 and extends to 2005, and a single realization is used for each model. The contour interval is 0.2°C in (a) and (b) and 0.1°C in (c), with the zero contour omitted. Stippling in (c) indicates where the differences are statistically significant at the 95% significance level based on a two-sided t-test. (d) Standard deviation of the PDO SST regressions within the ensemble. Contour interval is 0.05°C.

**Figure 18.** December-February PDO SAT and precipitation regression patterns over North America. Regressions of DJF SAT (a,c) and precipitation (b,d) on the PDO index in (a,b) observations and (c,d) the CMIP5 ensemble. The differences between the regression patterns (CMIP5 minus observations) are shown in (e) and (f). The contour interval is 0.1°C for the SAT regressions (a, c, e) and 0.05 mm/day for the precipitation regressions (b,d,f). Stippling in (e) and (f) correspond to differences that are significantly different at the 95% confidence level based on a two-sided t-test. To focus on multidecadal variability a Butterworth 10-year low-pass filter is applied to each PDO index time series, which is then re-standardized and detrended. The SAT and precipitation anomalies are then regressed on the filtered index for each season. The observations are the CRU TS3.1 temperature and precipitation datasets.

**Figure 19.** The JJASON AMO index in CMIP5 models compared to observations for (a) the time series and (b) autocorrelations. The AMO index is defined as the the 11-year running mean of the detrended North Atlantic SST during the Atlantic hurricane season of June to November (JJASON) from the equator to 60°N, 75°W-5°W. SST observations are from the ERSST dataset.

**Figure 20.** Autumn (SON) regressions of the AMO index on SST and precipitation from observations (HadISSTv1.1 and CRU TS3.1), and 17 CMIP5 models for 1901-1999. The AMO index is the area-averaged SST anomalies over the domain (75°-5°W, 0°-60°N), which are detrended and then smoothed via a 11-year running mean. Regressions are calculated for the first ensemble member for each model; observed and simulated anomalies have been regridded to a 1.5°×1.5° grid for precipitation, and a 5°×2.5° grid for SST. Blue/red shading denotes negative/positive SST anomalies, while brown/green shading denotes negative/positive precipitation anomalies. Contour interval is 0.1K and 0.02 mm day<sup>-1</sup>, respectively.

**Figure 21.** Observed and MME mean temperature trends (°C/decade) for North America (1930-2004) for (a) annual, (b) summer, and (c) winter. Observations are from the CRU TS3.1 dataset. The MME mean is calculated from the first ensemble member of 17 models (BCC-CSM1.1, CanESM2, CCSM4, CNRM-CM5, CSIRO-Mk3-6-0, GFDL-CM3, GFDL-ESM2M, GISS-E2-R, HadCM3, HadGEM2-ES, INMCM4, IPSL-CM5A-LR, MIROC5, MIROC-ESM, MPI-ESM-LR, MRI-CGCM3, NorESM1-M). Eastern and western US regions are shown by the boxes.

**Figure 22:** 30-year running annual temperature trend for (a) the Eastern US, (b) difference in trend between the Eastern and Western US. Regions are defined in Fig. 21. Shading represents the 95% uncertainty range calculated from 17 models (see Fig. 21), one ensemble member from each model. Black solid line is the observation (CRU TS3.1) and blue solid line is the MME median. X-axis represents the start of the 30-year running period. For example, the trend value at 1930 represents the trend from 1930 to 1959.

**Figure 23.** Comparison of (a) mean DTR and (b) DTR trend for the observations (CRU TS3.1) and MME mean of the 17 models (right; see Fig. 21) for 1951-2000. The contour lines are the inter-model standard deviation.

**Figure 24.** Observed and MME mean annual precipitation trends (mm/decade) for North America (1930-2004). Observations are from the CRU TS3.1 dataset. The MME mean is from 17 models (see Fig. 21), 1 ensemble member from each model. Eastern and western US regions are shown by the boxes.

**Figure 25.** 30-year running annual precipitation trend (mm/decade) for (a) the Eastern US, and (b) the Western US. Regions are shown in Fig. 23. The shaded region is the 95% uncertainty range calculated from 17 models (see Fig. 21), one ensemble member from each model. Black solid line is the observation (CRU TS3.1) and blue solid line is the MME median. X-axis represents the start of the 30-year running period. For example, the trend value at 1930 represents the trend from 1930 to 1959.



1622 **Tables**

1623 **Table 1.** CMIP5 models evaluated and their attributes.

Model	Center	Atmospheric Horizontal Resolution (lon. x lat.)	Number of model levels	Reference
ACCESS1-0	Commonwealth Scientific and Industrial Research Organization/Bureau of Meteorology, Australia	1.875 x 1.25	38	Bi et al. (2012)
BCC-CSM1.1	Beijing Climate Center, China Meteorological Administration, China	2.8 x 2.8	26	Xin et al. (2012)
CanCM4	Canadian Centre for Climate Modelling and Analysis, Canada	2.8 x 2.8	35	Merryfield et al. (2012)
CanESM2	Canadian Center for Climate Modeling and Analysis, Canada	2.8 x 2.8	35	Arora et al. (2011)
CCSM4	National Center for Atmospheric Research, USA	1.25 x 0.94	26	Gent et al. (2011)
CESM1-CAM5-1-FV2	Community Earth System Model Contributors (NSF-DOE-NCAR)	1.4 x 1.4	26	Gent et al. (2011)
CNRM-CM5.1	National Centre for Meteorological Research, France	1.4 x 1.4	31	Voldoire et al. (2011)
CSIRO-MK3.6	Commonwealth Scientific and Industrial Research Organization/Queensland Climate Change Centre of Excellence, AUS	1.8 x 1.8	18	Rotstayn et al. (2010)
EC-EARTH	EC-EARTH consortium	1.125 x 1.12	62	Hazeleger et al. (2010)
FGOALS-S2.0	LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences	2.8 x 1.6	26	Bao et al. (2012)

GFDL-CM3	NOAA Geophysical Fluid Dynamics Laboratory, USA	2.5 x 2.0	48	Donner et al. (2011)
GFDL-ESM2G/M	NOAA Geophysical Fluid Dynamics Laboratory, USA	2.5 x 2.0	48	Donner et al. (2011)
GISS-E2-H/R	NASA Goddard Institute for Space Studies, USA	2.5 x 2.0	40	Kim et al. (2012)
HadCM3	Met Office Hadley Centre, UK	3.75 x 2.5	19	Collins et al. (2001)
HADGEM2-CC (Chemistry coupled)	Met Office Hadley Centre, UK	1.875 x 1.25	60	Jones et al. (2011)
HadGEM2-ES	Met Office Hadley Centre, UK	1.875 x 1.25	60	Jones et al. (2011)
INMCM4	Institute for Numerical Mathematics, Russia	2 x 1.5	21	Volodin et al. (2010)
IPSL-CM5A-LR	Institut Pierre Simon Laplace, France	3.75 x 1.8	39	Dufresne et al. (2012)
IPSL-CM5A-MR	Institut Pierre Simon Laplace, France	2.5 x 1.25	39	Dufresne et al. (2012)
MIROC4h	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology, Japan	0.56 x 0.56	56	Sakamoto et al. (2012)
MIROC5	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology, Japan	1.4 x 1.4	40	Watanabe et al. (2010)
MIROC-ESM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The	2.8 x 2.8	80	Watanabe et al. (2010)

	University of Tokyo), and National Institute for Environmental Studies			
MIROC- ESM-CHEM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	2.8 x 2.8	80	Watanabe et al. (2010)
MPI-ESM- LR	Max Planck Institute for Meteorology, Germany	1.9 x 1.9	47	Zanchettin et al. (2012)
MRI-CGCM3	Meteorological Research Institute, Japan	1.1 x 1.1	48	Yukimoto et al. (2011)
NorESM1-M	Norwegian Climate Center, Norway	2.5 x 1.9	26	Zhang et al. (2012)

1624

1625

1625 **Table 2.** Observational and reanalysis datasets used in the evaluations.

Dataset	Type	Spatial Domain	Temporal Domain	Reference
<i>Precipitation</i>				
TMPA 3B42 V6	Satellite	0.25 deg, 50S-50N	3-hourly/monthly, 1998-2010	Huffman et al. (2007)
CRU TS3.1	Gauge	0.5 deg, global land	Monthly, 1901-2008	Mitchell et al. (2005)
CPC unified	Gauge	0.5 deg, US	Daily, 1948-2010	Xie et al. (2010)
GPCP v2.1	Gauge/satellite	1.0 deg, global	1979-2009	Adler et al. (2003)
UNAM v0705	Gauge	0.5 deg, Mexico and surroundings	1901-2002	UNAM (2007)
<i>Temperature</i>				
CRU TS3.1	Gauge	0.5 deg, global land	Monthly, 1901-2008	Mitchell et al. (2005)
<i>Sea Surface Temperature</i>				
HadISST	In situ/satellite	Global Oceans, 1.0 deg	Monthly, 1870-present	Rayner et al. (2003)
ERSSTv3b	In situ	Global Oceans, 2.0 deg	Monthly, 1854-present	Smith et al. (2008)
<i>Reanalyses</i>				
NCEP-NCAR	Model reanalysis	~1.9 deg, global	6-hourly, 1948-present	Kalnay et al. (1996)
NCEP-DOE	Model reanalysis	~1.9 deg, global	6-hourly, 1979-present	Kanamitsu et al. (2002)
20CR	Model reanalysis	~0.3 deg, global	6-hourly, 1871-present	Compo et al. (2011)
ERA-interim	Model reanalysis	1.5 deg, global	6-hourly, 1979-present	Dee et al. (2011)
<i>Storm Tracks</i>				
National Hurricane Center best track tropical cyclone data	In-situ/satellite	Storm tracks, Eastern N. Pacific and N. Atlantic	6-hourly, 1851-present (Atlantic), 1949-present (eastern N. Pacific)	Jarvinen et al. (1984)
IBTrACS	Best track datasets from various agencies	Storm tracks, global oceans	6-hourly, 1842-2010	Knapp et al. (2010)

1626

1627

1627 **Table 3.** Spatial correlation of the MSD between the CMIP5 models and the MME mean,  
1628 calculated for 1850-2005 .

Model	R
BCC-CSM1-1	0.45
CanCM4	0.37
CanESM2	0.42
CCSM4	0.17
CNRM-CM5	0.49
CSIRO-Mk3-6-0	0.51
GFDL-CM3	0.29
GFDL-ESM2G	0.48
GFDL-ESM2M	0.27
GISS-E2-H	0.35
GISS-E2-R	0.34
HadCM3	0.75
HadGEM2-CC	0.79
HadGEM2-ES	0.81
INMCM4	0.14
IPSL-CM5A-LR	0.40
IPSL-CM5A-MR	0.34
MIROC5	0.71
MIROC-ESM	-0.04
MIROC-ESM-CHEM	-0.04
MPI-ESM-LR	0.61
MRI-CGCM3	0.33
NorESM1-M	0.14

1629

1630



**Table 4.** Spatial correlations of model fields with ERA-Interim for the months indicated and for 1979-2005. Correlations of the 850 hPa wind components and geopotential height have been combined into one index R\_ZUV850, while 850 hPa track density and strength correlations have been combined into a second index R\_TRK850 to simplify the comparisons. Values in bold are the upper 25<sup>th</sup> percentile of the nine models shown.

	May-Oct R_ZUV850	May-Nov R_TRK850
BCC-CSM1.1	0.76	0.69
CanESM2	<b>0.83</b>	0.63
CCSM4	0.77	0.57
CNRM-CM5	<b>0.90</b>	<b>0.84</b>
GFDL-ESM2M	0.75	0.77
HadGEM2-ES	<b>0.85</b>	<b>0.90</b>
MIROC5	0.82	0.86
MPI-ESM-LR	0.82	0.85
MRI-CGCM3	0.79	<b>0.86</b>
75 <sup>th</sup> Percentile	0.83	0.86

**Table 5.** Error statistics for the CMIP5 PDO regressions on North American seasonal SAT and temperature. Pattern correlations lie above the RMS difference (°C for SAT, mm/day for precipitation). Regression differences at each grid point are evaluated with a two-sided t-test, for which the effective degrees of freedom are adjusted for the lag-1 autocorrelation in the residuals, as in Santer et al. (2000). Bold indicates regression pattern differences that are statistically significant at the 5% level based on a false discovery rate field significance test (Wilks, 2006).

Model	DJF SAT	DJF Precip.	MAM SAT	MAM Precip.	JJA SAT	JJA Precip.	SON SAT	SON Precip.
BCC-CSM1.1	0.90 0.18	0.33 0.07	0.67 0.12	0.20 0.05	0.48 0.08	0.07 0.10	0.05 0.16	0.06 0.07
CanESM2	0.63 0.19	0.31 0.06	0.76 0.10	0.24 0.05	0.45 0.07	0.04 0.06	-0.11 0.12	-0.07 0.07
CCSM4	0.85 0.16	0.44 0.06	0.69 0.20	0.23 0.06	0.29 0.09	-0.15 0.08	<b>0.10</b> <b>0.29</b>	<b>0.11</b> <b>0.10</b>
CNRM-CM5	0.54 0.22	0.30 0.07	0.64 0.18	0.20 0.06	0.31 0.11	-0.04 0.07	0.44 0.09	-0.02 0.08
CSIRO-Mk3.6.0	0.70 0.19	0.11 0.07	0.51 0.20	<b>-0.05</b> <b>0.11</b>	0.31 0.14	-0.01 0.19	0.30 0.18	0.16 0.12
FGOALS-s2	0.72 0.16	0.02 0.08	0.41 0.15	0.09 0.06	0.26 0.10	0.00 0.06	0.34 0.14	-0.08 0.07
GFDL-CM3	0.53 0.20	0.24 0.07	0.19 0.18	-0.11 0.06	0.22 0.10	0.01 0.06	0.19 0.14	0.10 0.06
GFDL-ESM-2G	0.82 0.25	<b>0.06</b> <b>0.07</b>	0.53 0.23	-0.01 0.07	0.28 0.12	-0.06 0.09	0.28 0.18	0.14 0.09
GFDL-ESM-2M	0.52 0.30	0.26 0.08	<b>0.22</b> <b>0.22</b>	<b>-0.17</b> <b>0.08</b>	0.35 0.11	0.03 0.11	0.01 0.14	0.07 0.09
GISS-E2-R	0.70 0.17	0.41 0.06	0.57 0.16	0.18 0.06	0.04 0.12	-0.04 0.10	0.11 0.11	-0.04 0.08
HadGEM2-CC	0.76 0.16	0.52 0.06	0.55 0.20	0.26 0.06	0.37 0.11	-0.03 0.08	0.42 0.09	0.19 0.06
HadGEM2-ES	0.39 0.22	0.20 0.08	0.62 0.19	0.10 0.07	0.26 0.09	-0.09 0.11	0.07 0.12	0.08 0.09
HadCM3	0.73 0.20	0.30 0.07	0.75 0.13	0.28 0.05	0.21 0.11	0.04 0.08	0.27 0.14	0.20 0.08
INM-CM4	0.15 0.25	0.13 0.07	0.14 0.19	-0.14 0.07	0.05 0.10	-0.02 0.07	0.20 0.09	0.05 0.06
IPSL-CM5A-LR	0.87 0.15	0.21 0.09	<b>0.40</b> <b>0.23</b>	-0.13 0.08	<b>0.14</b> <b>0.10</b>	<b>-0.14</b> <b>0.07</b>	0.19 0.13	0.09 0.07
IPSL-CM5A-MR	0.74 0.17	0.18 0.09	0.42 0.17	-0.09 0.07	0.11 0.10	-0.10 0.07	0.42 0.11	-0.06 0.07
MIROC5	0.65 0.19	0.23 0.11	0.69 0.14	0.26 0.06	0.29 0.10	0.00 0.08	0.05 0.15	-0.02 0.09
MIROC-ESM	0.48 0.21	-0.07 0.07	<b>0.13</b> <b>0.17</b>	-0.22 0.06	0.33 0.07	-0.11 0.06	0.31 0.09	0.05 0.06
MPI-ESM-LR	0.77 0.18	0.20 0.07	0.34 0.18	-0.03 0.07	0.00 0.13	0.00 0.08	0.20 0.14	0.10 0.08
MRI-CGCM3	-0.47 0.34	0.11 0.08	0.09 0.24	0.10 0.08	0.14 0.08	-0.04 0.08	-0.64 0.21	-0.04 0.08
NorESM1-M	0.76 0.24	0.14 0.08	0.52 0.17	-0.01 0.06	0.39 0.09	-0.07 0.07	-0.05 0.16	-0.09 0.07
MME mean	0.91 0.11	0.47 0.06	0.63 0.13	<b>0.10</b> <b>0.05</b>	0.37 0.07	-0.06 0.06	0.22 0.10	0.14 0.06

1645 **Table 6.** The RMSE and standard deviations of the AMO indices in CMIP5 models.  
1646 Observations are from the ERSST dataset.  
1647

Model Name	RMSE (°C)	Standard Deviation (°C)
ACCESS1-0	0.1846	0.1870
BCC-CSM1-1	0.1052	0.1528
CanESM2	0.1532	0.1442
CCSM4	0.1438	0.1198
CNRM-CM5	0.1529	0.1031
CSIRO-Mk3-6-0	0.1609	0.1550
EC-EARTH	0.1501	0.0914
FGOALS-g2	0.1835	0.1083
GFDL-CM3	0.1638	0.1598
GFDL-ESM2G	0.2110	0.1699
GFDL-ESM2M	0.1493	0.1273
GISS-E2-H	0.1376	0.0958
GISS-E2-R	0.1453	0.1054
HadCM3	0.1662	0.1421
HadGEM2-CC	0.1926	0.1895
HadGEM2-ES	0.1455	0.1517
INMCM4	0.1485	0.0917
IPSL-CM5A-LR	0.1800	0.1760
IPSL-CM5A-MR	0.1374	0.1320
IPSL-CM5B-LR	0.2240	0.1879
MIROC5	0.1347	0.1335
MIROC-ESM	0.1375	0.1467
MIROC-ESM-CHEM	0.1544	0.1364
MPI-ESM-LR	0.2123	0.1794
MPI-ESM-P	0.1526	0.0993
MRI-CGCM3	0.1515	0.1234
NorESM1-M	0.1366	0.1118
MME mean	0.1598	0.1378
Observations	0	0.1761

**Table 7.** Spatial correlation between observed and CMIP5 regressed anomalies of the AMO on SST and precipitation in summer and fall for 1901-1999. The spatial domain for SST correlations is over the Atlantic Ocean north of the equator (130°W-10°E, 0°-75°N), while the domain for precipitation is the American continent north of the equator (130°W-60°W, 0°-60°N). The observed data sets are the HadISSTv1.1 and CRUTS3.1 data sets for SST and precipitation, respectively.

Model	Summer SST	Fall SST	Summer Precipitation	Fall Precipitation
BCC-CSM1.1	-0.132	-0.205	0.131	0.293
CanESM2	0.459	0.597	0.080	-0.172
CCSM4	0.224	0.332	-0.092	-0.172
CNRM-CM5.1	0.527	0.037	-0.029	-0.357
CSIRO-MK3.6	0.037	0.308	-0.034	0.211
GFDL-CM3	-0.213	0.176	0.143	0.145
GFDL-ESM2M	0.325	0.461	0.129	0.014
GISS-E2-R	0.586	0.675	-0.070	-0.014
HadCM3	0.531	0.578	0.008	-0.116
HadGEM2-ES	0.700	0.485	0.172	-0.309
INM-CM4	-0.337	-0.126	-0.183	0.025
IPSL-CM5A-LR	0.180	0.327	-0.072	0.060
MIROC5	0.433	0.588	-0.196	-0.002
MIROC-ESM	0.430	0.384	-0.168	-0.033
MPI-ESM-LR	-0.135	0.230	-0.149	-0.129
MRI-CGCM3	0.412	0.215	0.335	0.140
NorESM1-M	0.098	-0.298	-0.127	-0.081
MME mean	0.577	0.651	-0.012	-0.033

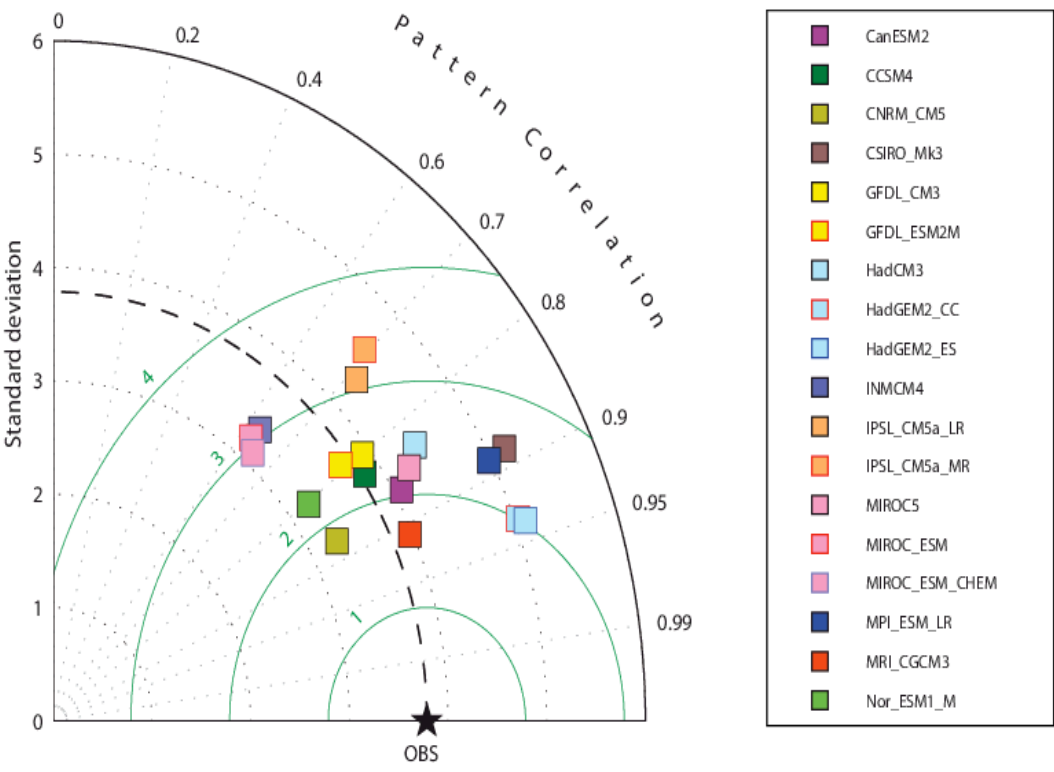
1648

1649

1650

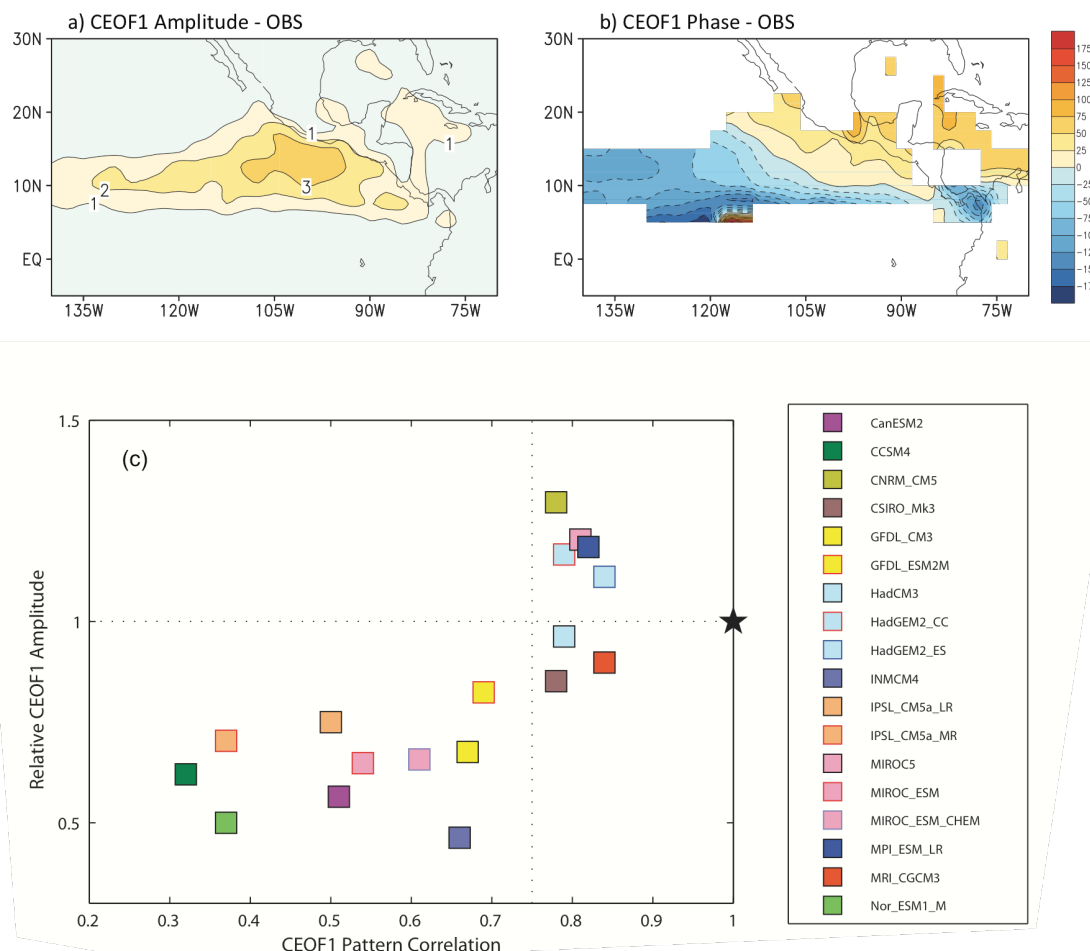
1651

1651

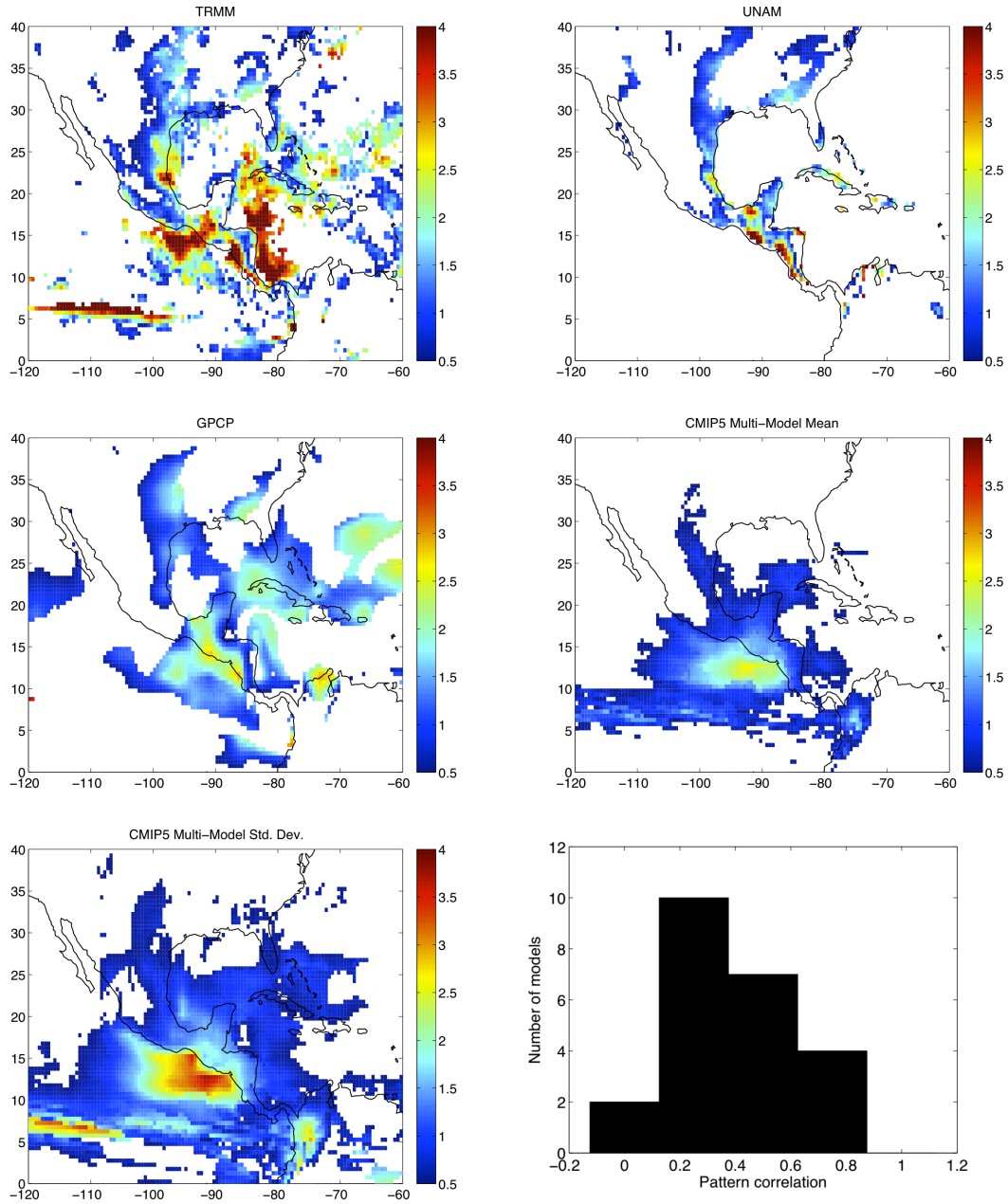


1652  
1653  
1654  
1655

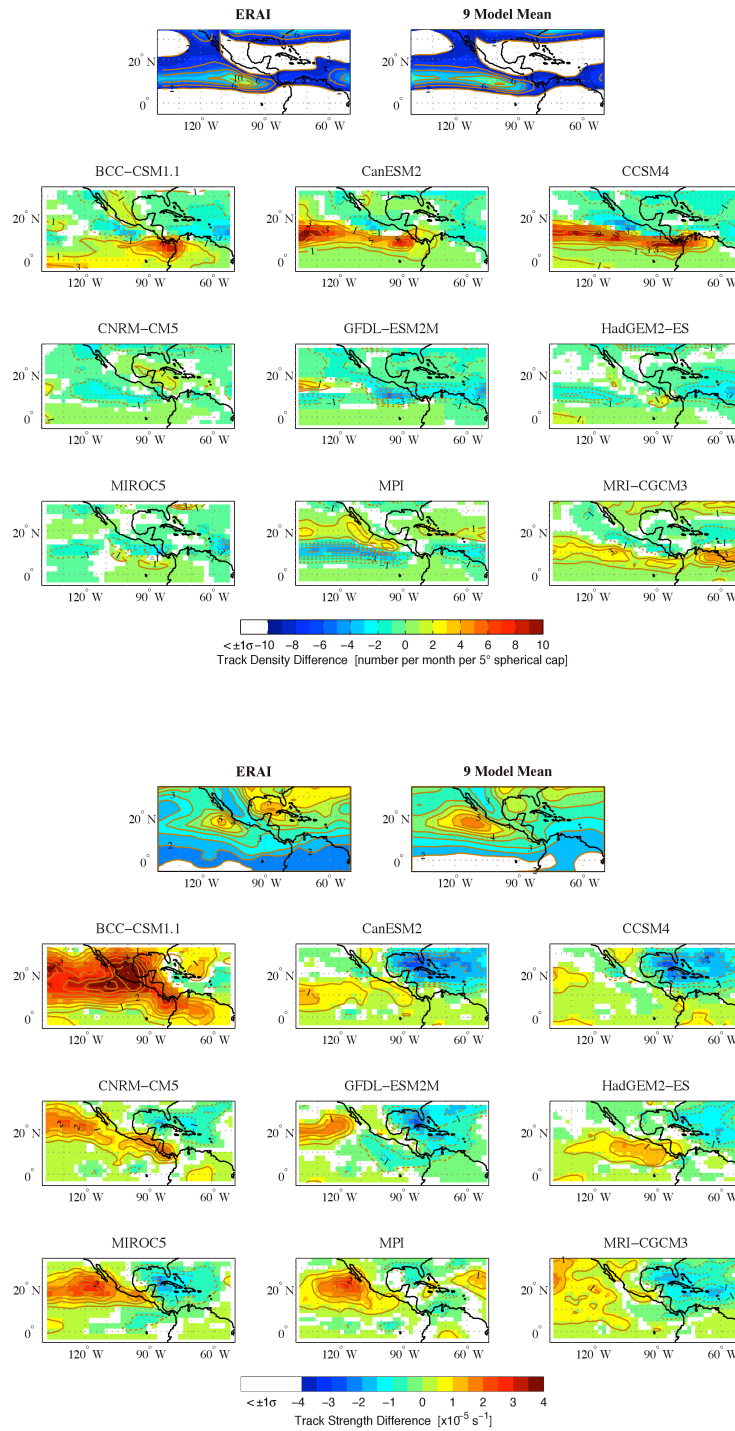
**Figure 1.** Taylor diagram for summer mean (May-September) rainfall over the eastern Pacific (150°W-80°W; 5°S-30°N) simulated in CMIP5 GCMs. The rainfall observations are based on TMPA data.



**Figure 2.** Spatial distribution of amplitude (a) and phase (b) of the first leading complex EOF (CEOF1) mode based on 30-90-day band-pass filtered TRMM rainfall during boreal summer (June-September) over the eastern Pacific. To make the spatial phase patterns of the CEOF1 based on the observations and simulations comparable to each other, the spatial phase of CEOF1 for each dataset is adjusted by setting the domain averaged value to be zero over a small box region of 110°W-100°W, 10-15°N. Contours are only displayed where the local variance explained by CEOF1 exceeds 8%; (c): X-axis: Pattern correlation coefficients of the CEOF1 mode between TRMM observations and CMIP5 GCM simulations. Y-axis: Relative amplitudes of CEOF1 in model simulations to their observed counterparts. Both pattern correlations and amplitudes are derived by averaging over the area of 5°N-25°N, 140°W-80°W where the active ISV is observed. The black “star” mark represents the TMPA observations.

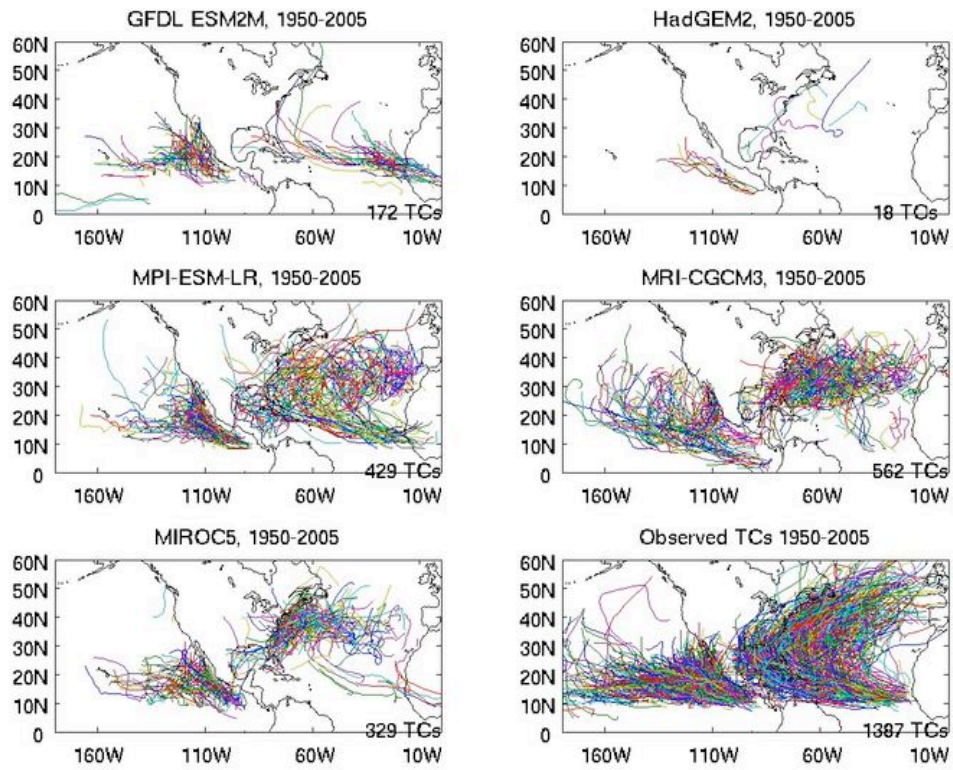


**Figure 3.** Summertime (June-September) MSD strength (mm/day) for three observational estimates (TRMM 3B43, UNAM and GPCP) and the CMIP5 MME mean for 23 models (see Table 3). Also shown (bottom row) are the MME standard deviation and histogram of the pattern correlations between individual models and the MME mean. All model output and observational data were regridded onto a common 0.5° grid.

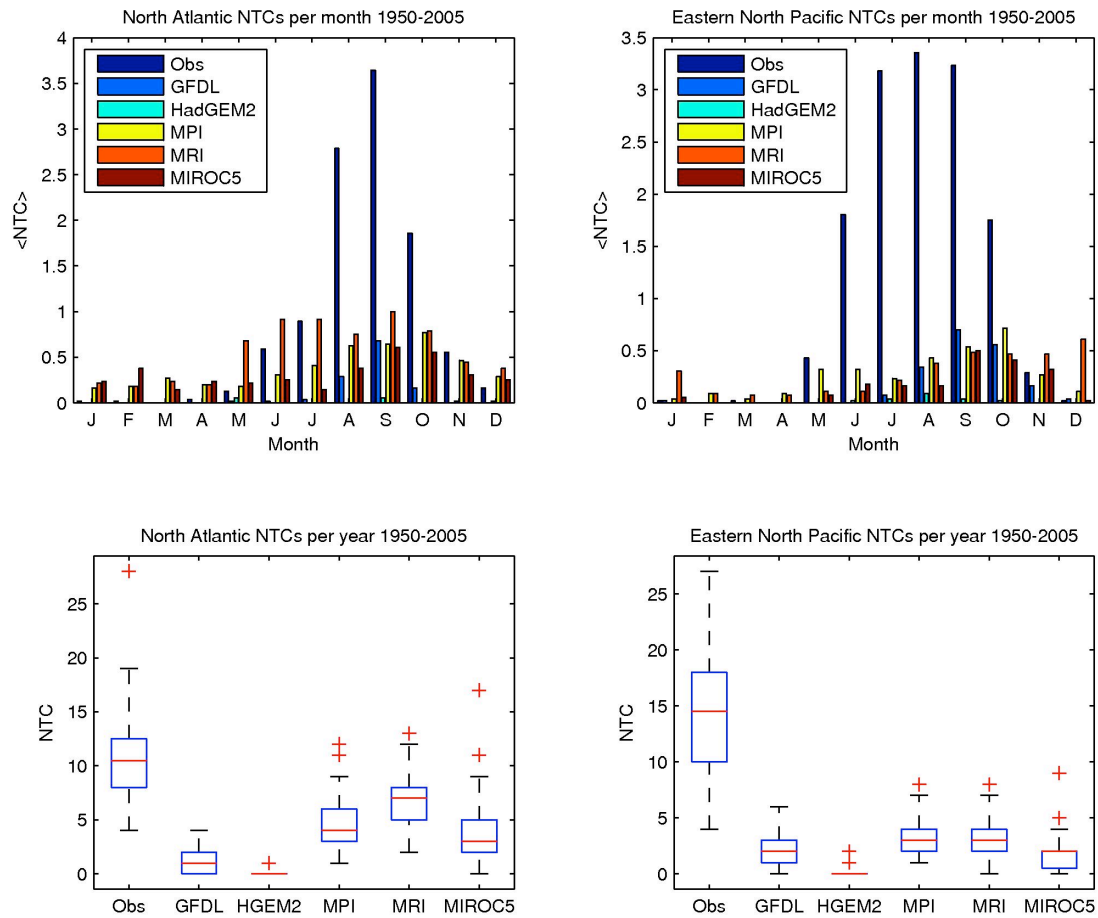


**Figure 4.** Storm track density (top) and mean strength (bottom) for ERA Interim and seven CMIP5 models (CanESM2, CCSM4, GFDL-ESM2M, HadGEM2-ES, MIROC5, MPI-ESM-L and MRI-CGCM3). Tracks are based on 6-hourly 850hPa relative vorticity smoothed to T42 spatial resolution to better capture the synoptic features of the vorticity field.

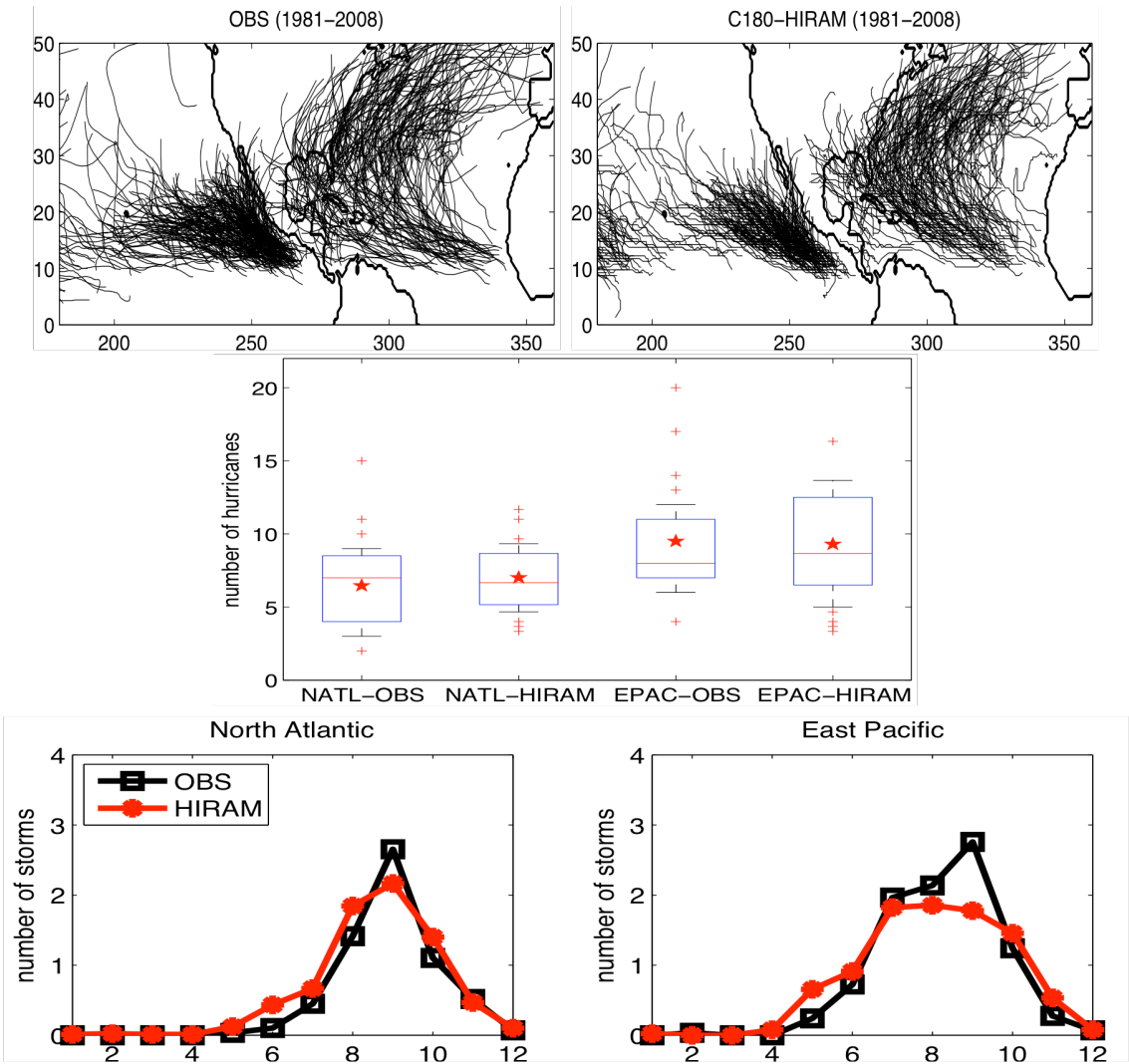




**Figure 5.** Tracks of tropical cyclone-like storms in the CMIP5 historical runs in the period 1950-2005 (GFDL-ESM2M (1 ensemble member), HadGEM2 (1), MPI-ESM-LR (3), MRI CGCM3 (5) and MIROC5 (1) models) and in observations for the same period. The number of storms in each case is given in the bottom right corner of each panel. One ensemble member is used for each model.



**Figure 6.** Mean number of TCs per month in models (GFDL-ESM2M, HadGEM2-ES (in the figure HGEM2), MPI-ESM-LR, MRI-CGCM3, MIROC5) and observations in the North Atlantic (top left panel) and eastern North Pacific (top right panel), using only ensemble 1 for MRI-CGCM3. Number of TCs per year in the period 1950-2005 in models and observations for the North Atlantic (bottom left panel) and eastern North Pacific (bottom right panel). The blue box shows the 25-75 percentile range, with the median shown as a red line. The whiskers and red crosses show the data outside of middle quartiles.



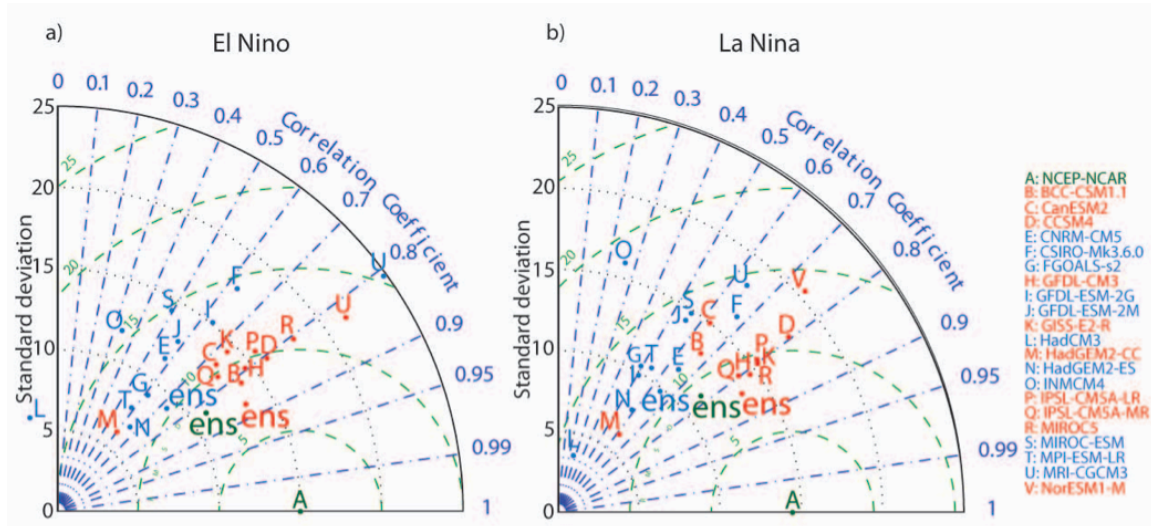
1700

1701 **Figure 7.** Upper panels: Comparison of observed and C180HIRAM (one realization)  
1702 simulated hurricane tracks for the N. Atlantic and E. Pacific for 1981-2008. Middle  
1703 panel: Comparison of observed and C180HIRAM simulated annual hurricane count  
1704 statistics. Blue boxes show the 25-75 percentile range, with the median shown as a red  
1705 line and the mean shown as a red star. The whiskers show the maximum and minimum  
1706 values. The annual statistics are computed based on a 3-member ensemble mean for  
1707 1981-2008. Lower panels: Observed and model simulated seasonal cycle (number of  
1708 hurricanes per month) for the N. Atlantic and E. Pacific from the 3-member ensemble  
1709 mean (1=JAN, 12=DEC)

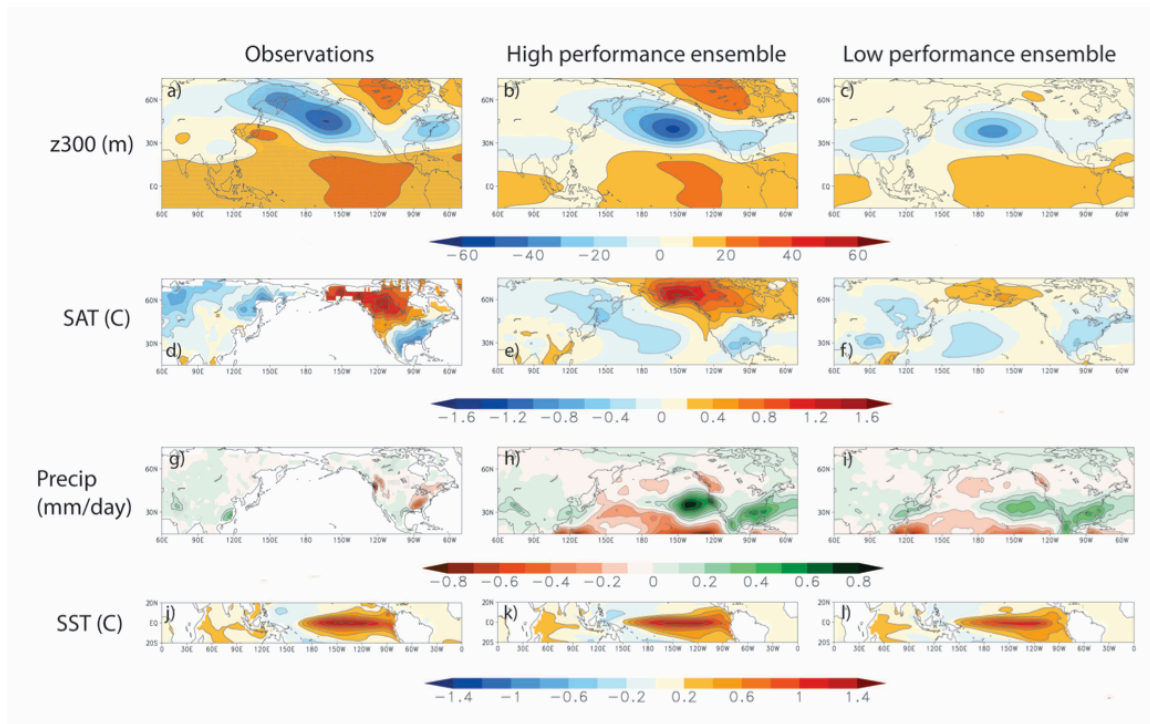
1710

1711

1712

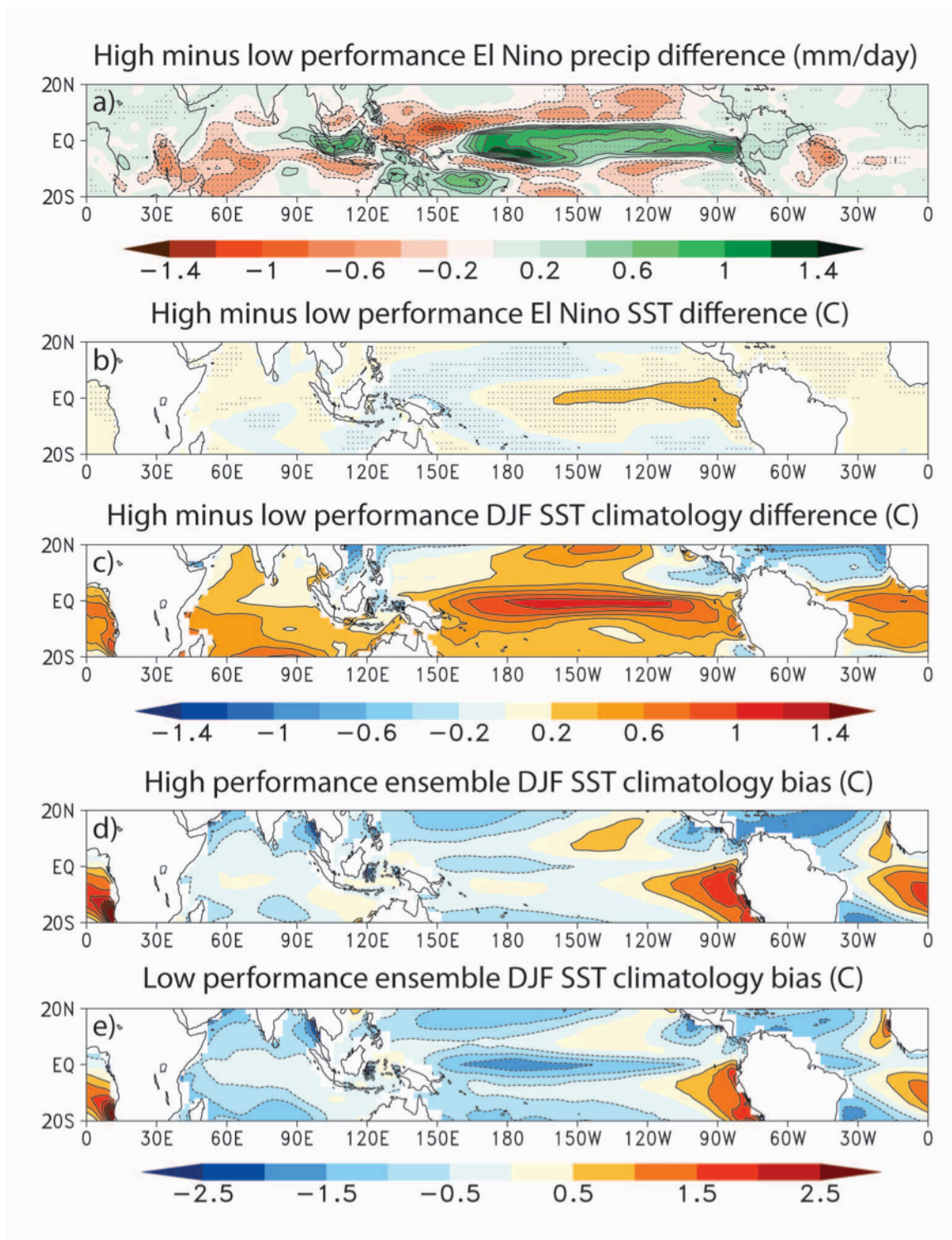


**Figure 8.** Taylor diagrams for (a) El Niño and (b) La Niña composite 300 hPa geopotential patterns over the region from East Asia – North America. Higher performance (pattern correlation > 0.6, RMS difference < 13m in both (a) and (b)) models are indicated in red, whereas lower performance models are indicated in blue. In (a) HadCM3, which falls outside of the plot, has a pattern correlation of -0.3 and RMS difference of 17.6 m. The points labeled “ens” in red, blue, and green represent the higher performance, lower performance, and total ensemble, respectively. The composites are normalized by the Niño 3.4 SST amplitude to focus on pattern differences independent of ENSO amplitude differences. The observational reference is based on the NCEP/NCAR reanalysis for 1950-2010, whereas the CMIP calculations are based on the full historical period (1850-2005) for one run of each model.

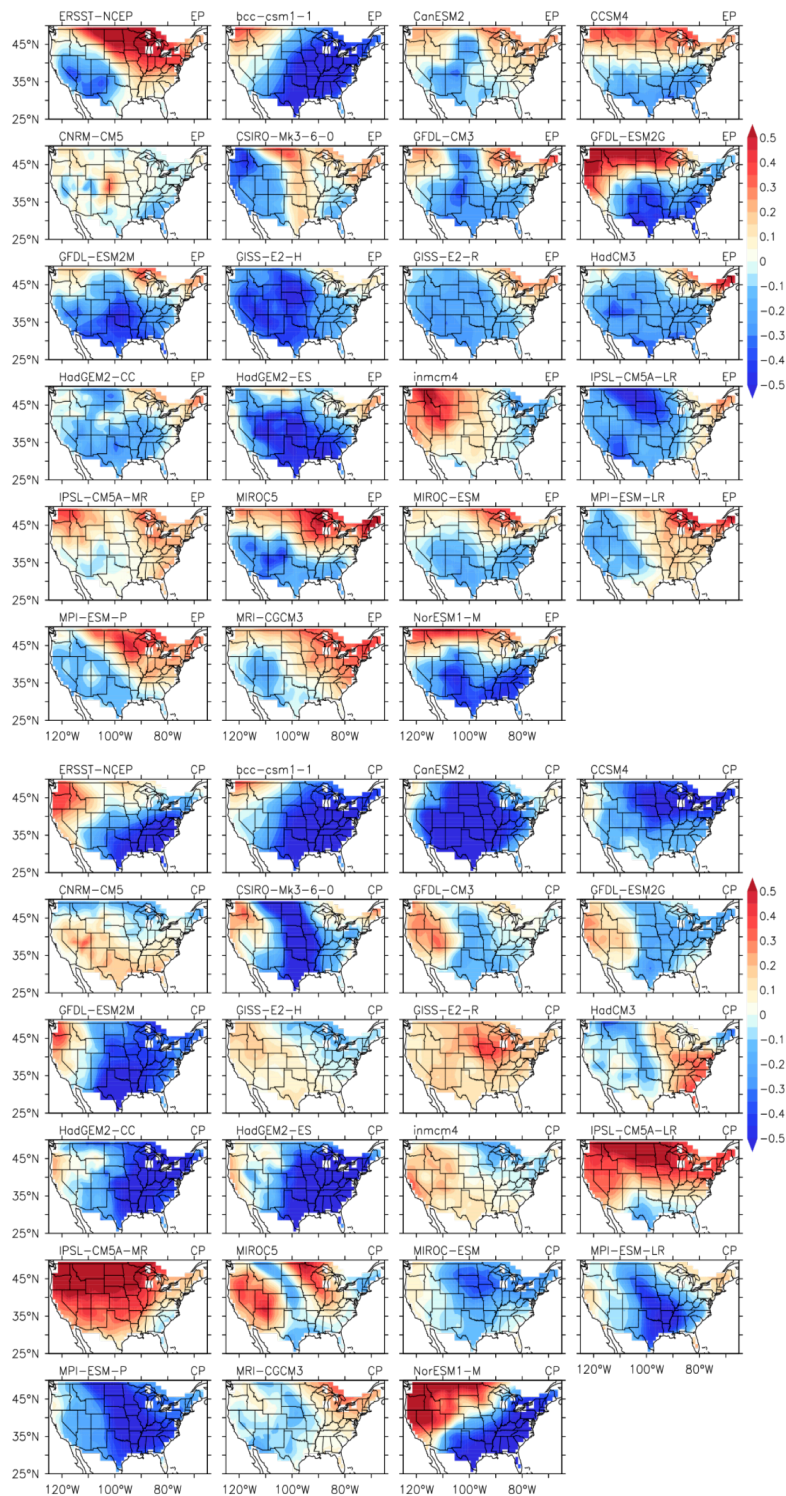


**Figure 9.** Composites of (a,b,c) 300 hPa height (z300, m), (d,e,f) near surface air temperature (SAT, °C), (g,h,i) precipitation (Precip, mm/day), and (j,k,l) SST (°C) anomalies during DJF El Niño episodes in observations (left column), and both high (middle column) and low performance CMIP5 ensembles (right column) described in Fig. 8. The observational SAT and precipitation composites are based on the CRU TS3.1 land near-surface temperature and precipitation datasets for 1901-2009. The z300, SAT, and Precip composites are normalized by the Niño 3.4 SST anomaly. Stippling in the observed (a) z300, (d) SAT, and (g) precipitation composites indicates anomalies that are statistically significant at the 5% level.

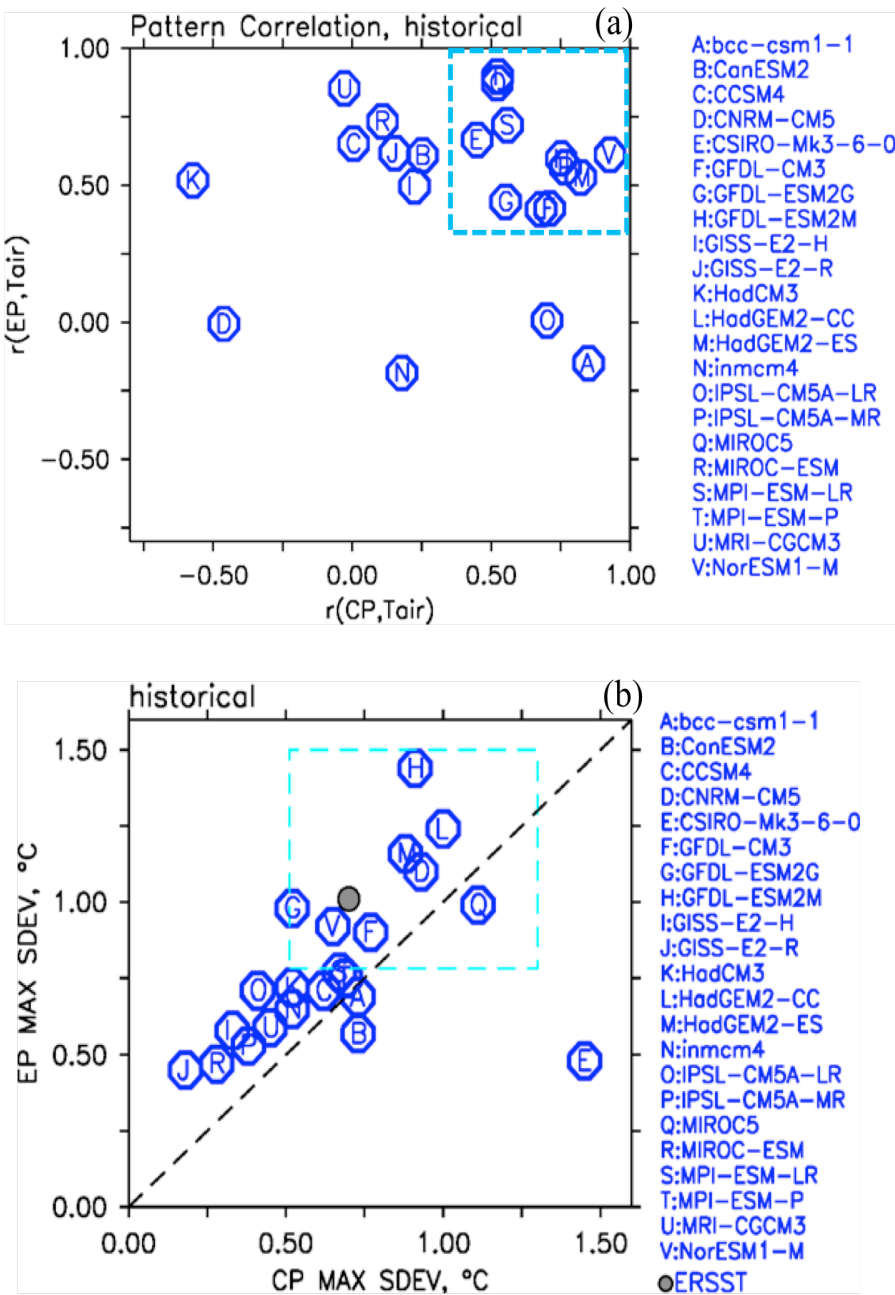




**Figure 10.** Composite DJF El Niño (a) precipitation (mm/day) and (b) SST (°C) difference between the high and low performance CMIP5 ensemble described in Fig. 8. Stippling indicates differences that are statistically significant at the 5% level. (c) DJF SST climatology difference (°C) between the high and low performance ensemble, and (d) high and (e) low performance SST climatology bias (°C) for the 1951-2000 period.

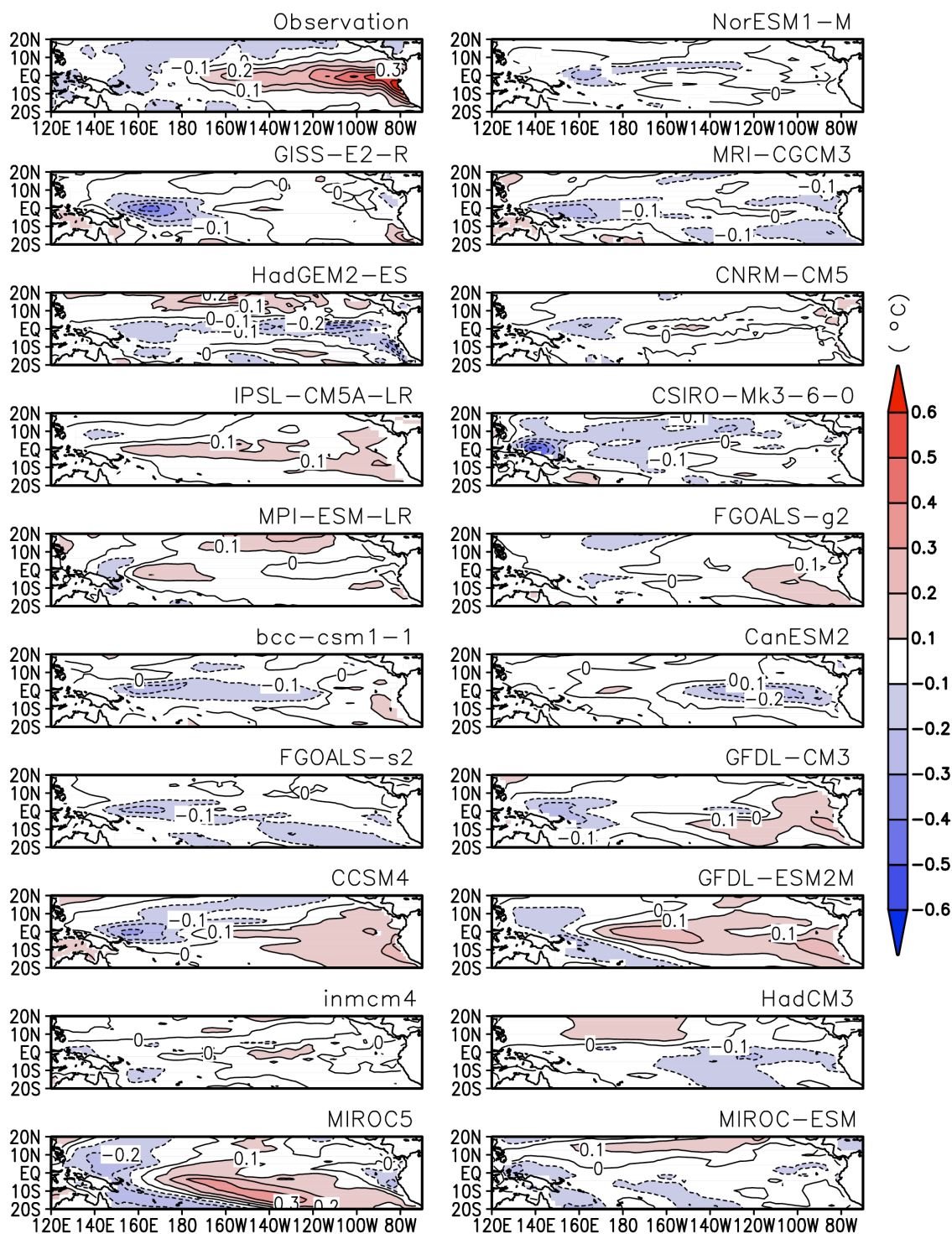


**Figure 11.** US winter surface air temperature regressed on the EP (top six rows) and CP (bottom six rows) ENSO indices from observations and the CMIP5 models. Observational air temperature data are from the NCEP-NCAR reanalysis and SSTs are from the ERSST dataset for 1950-2010.

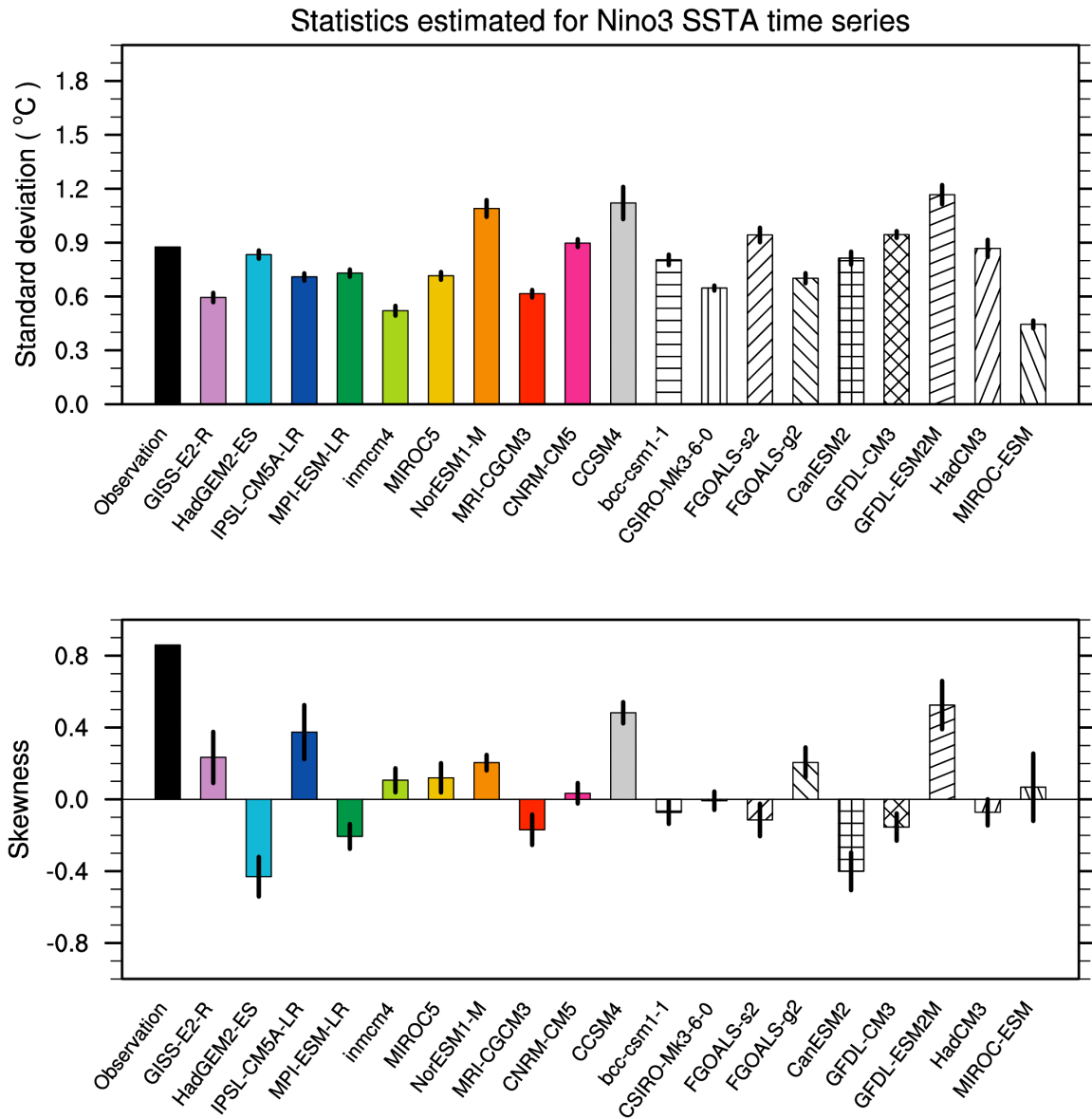


**Figure 12.** (a) Scatter plot of pattern correlations between the regression patterns from the CMIP5 models and those from the observations (NCEP-NCAR reanalysis and HadISST dataset) for EP versus CP ENSO; (b) Scatter plot of the intensities of the EP/CP ENSO from the CMIP5 models and the observation (ERSST). The values shown are the maximum standard deviations of the EOF patterns of the two types of the ENSO calculated using a regression-EOF method. The blue dashed lines indicate the lower limit of the 95% significance interval of the observed ENSO intensities based on an F-test.

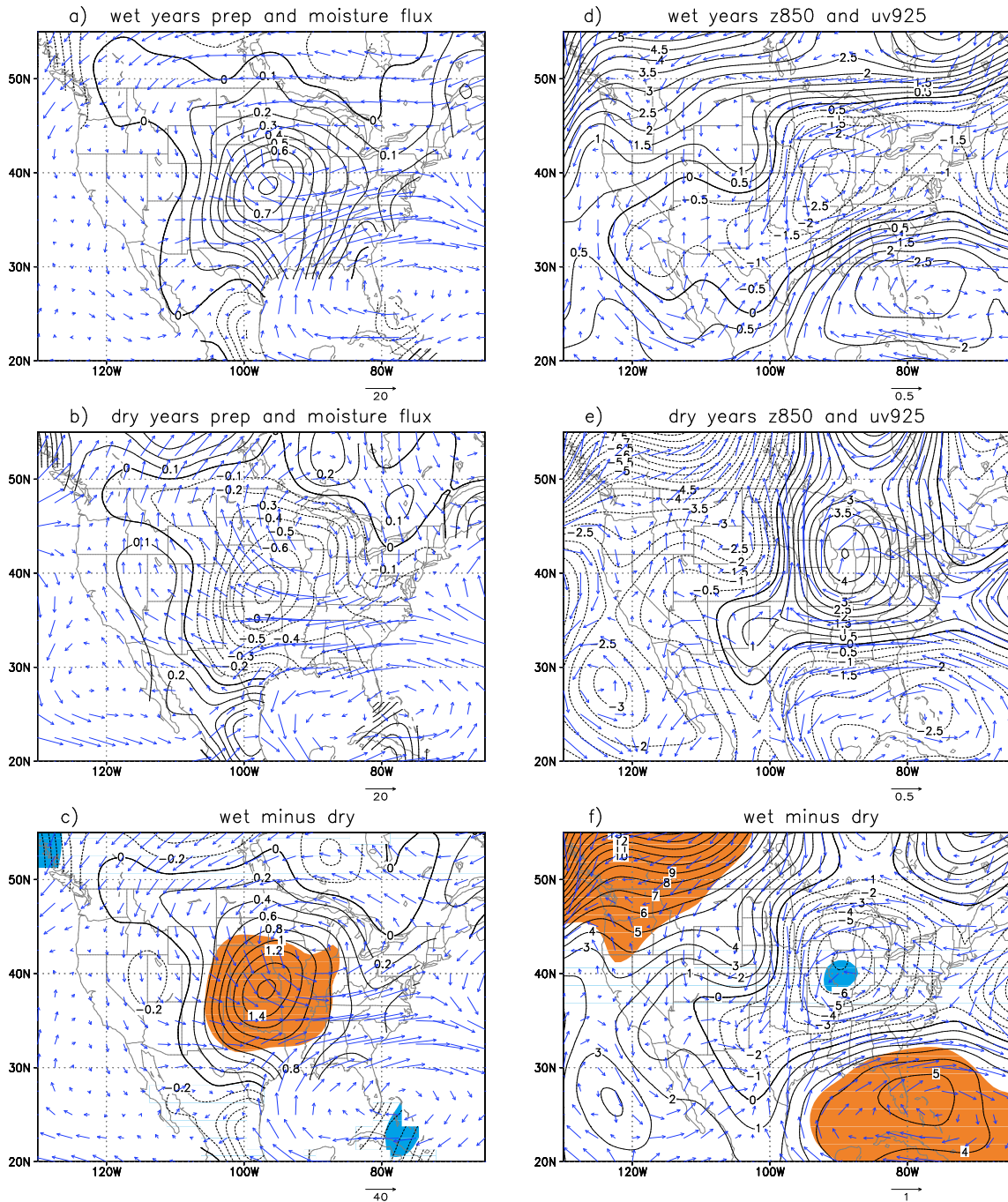




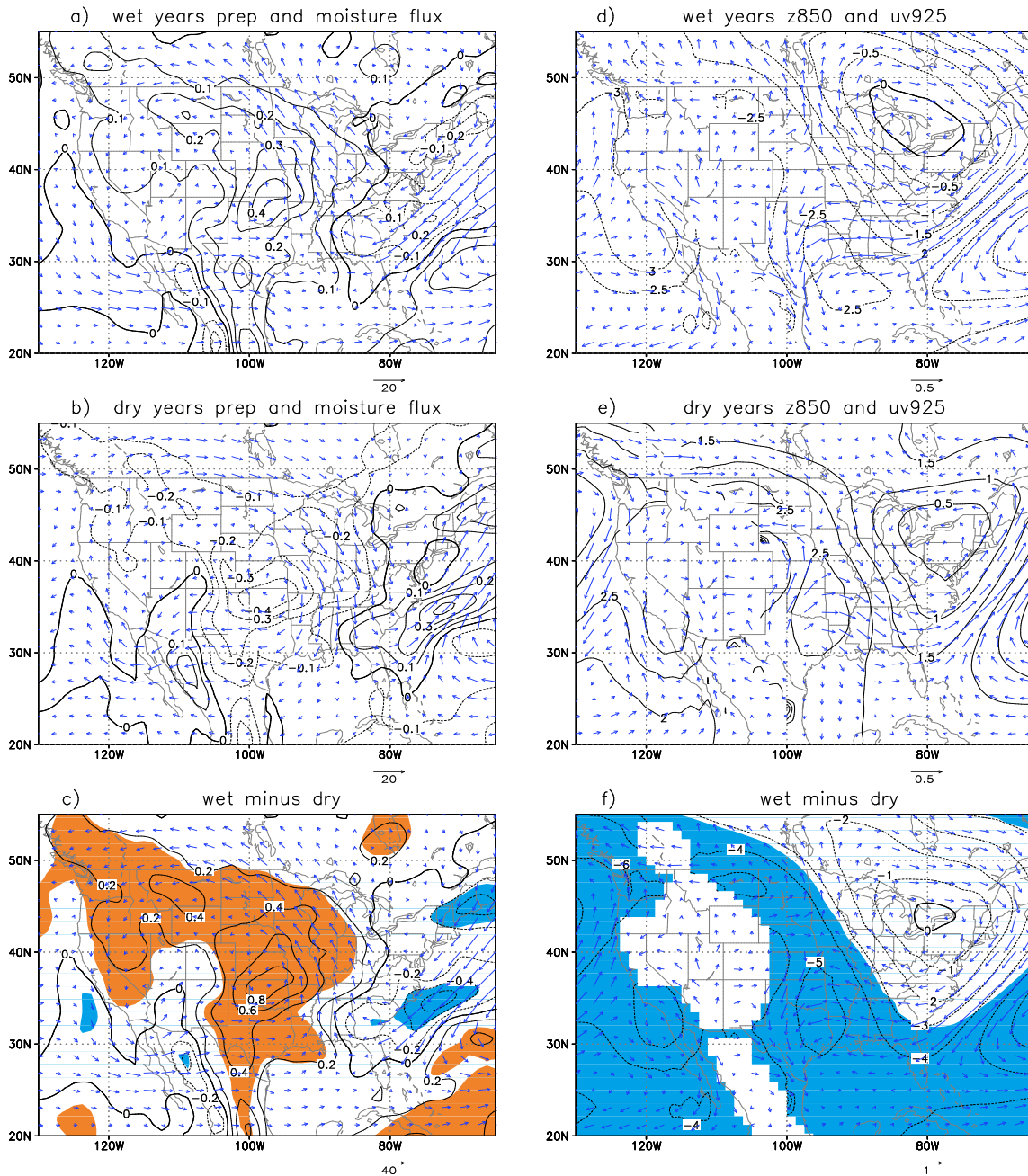
**Figure 13.** The sum of the composite SST anomalies between the two phases of ENSO from the HadISST observations and CMIP5 coupled models. The definition of the warm phase and cold phase of ENSO follows that of Zhang et al. (2009). The length of data used in the calculation is 50 years for all the models and observations (1950-99).



**Figure 14.** The standard deviation (upper panel) and skewness (lower panel) of monthly Niño-3 SST anomalies from observations and CMIP5 model simulations. The length of data for computing the standard deviation and skewness is 50 years for the observations (1950–99). For the model, the standard deviation and skewness were calculated for a 50-year moving window over 100 years of the model run for a total of 601 samples. The figure shows the mean of the samples and +/- the standard deviation across the samples. Data used are the same as for Figure 13.

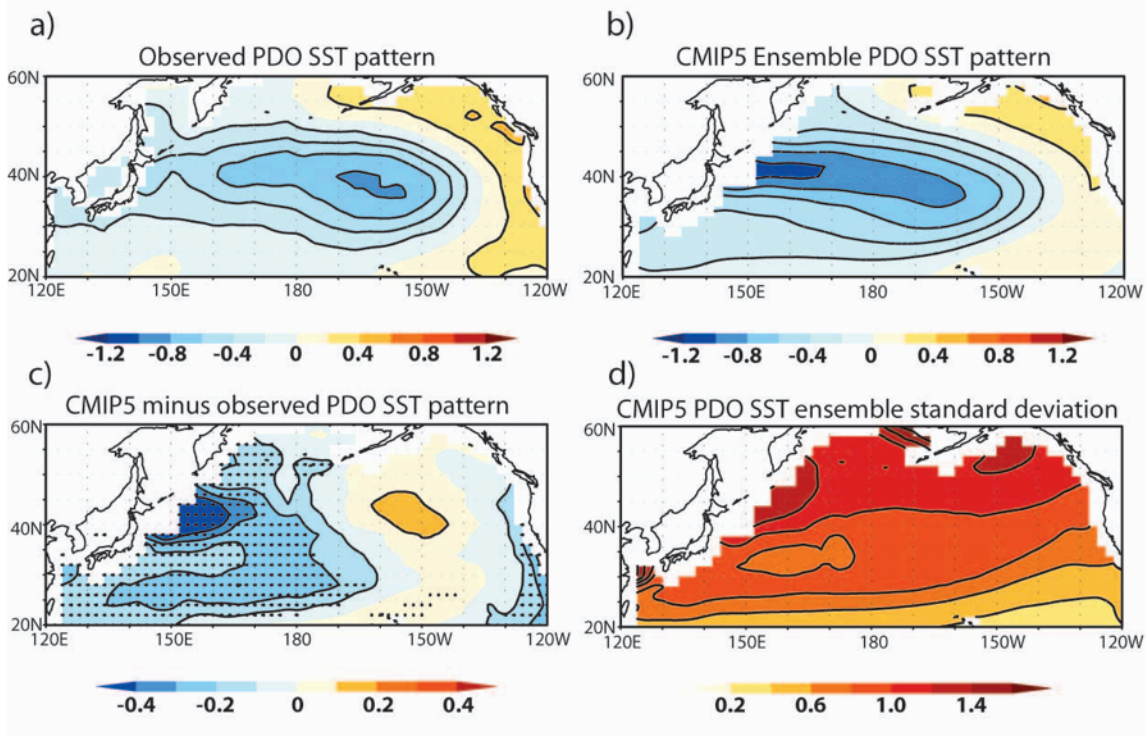


**Figure 15.** Summertime wet and dry circulation patterns for the central US from the NCEP/NCAR reanalysis. a) and b) show, respectively, summertime precipitation anomalies (contours) in wet and dry years, in reference to the Great Plains precipitation, and the vertically integrated moisture fluxes from the surface to the top of the troposphere (arrows). c) The differences between a) and b). d) and e) show the corresponding 850hPa geopotential height (countour) and 925hPa wind anomalies (arrows) for the wet and dry summer, respectively. Their differences are summerized in f).

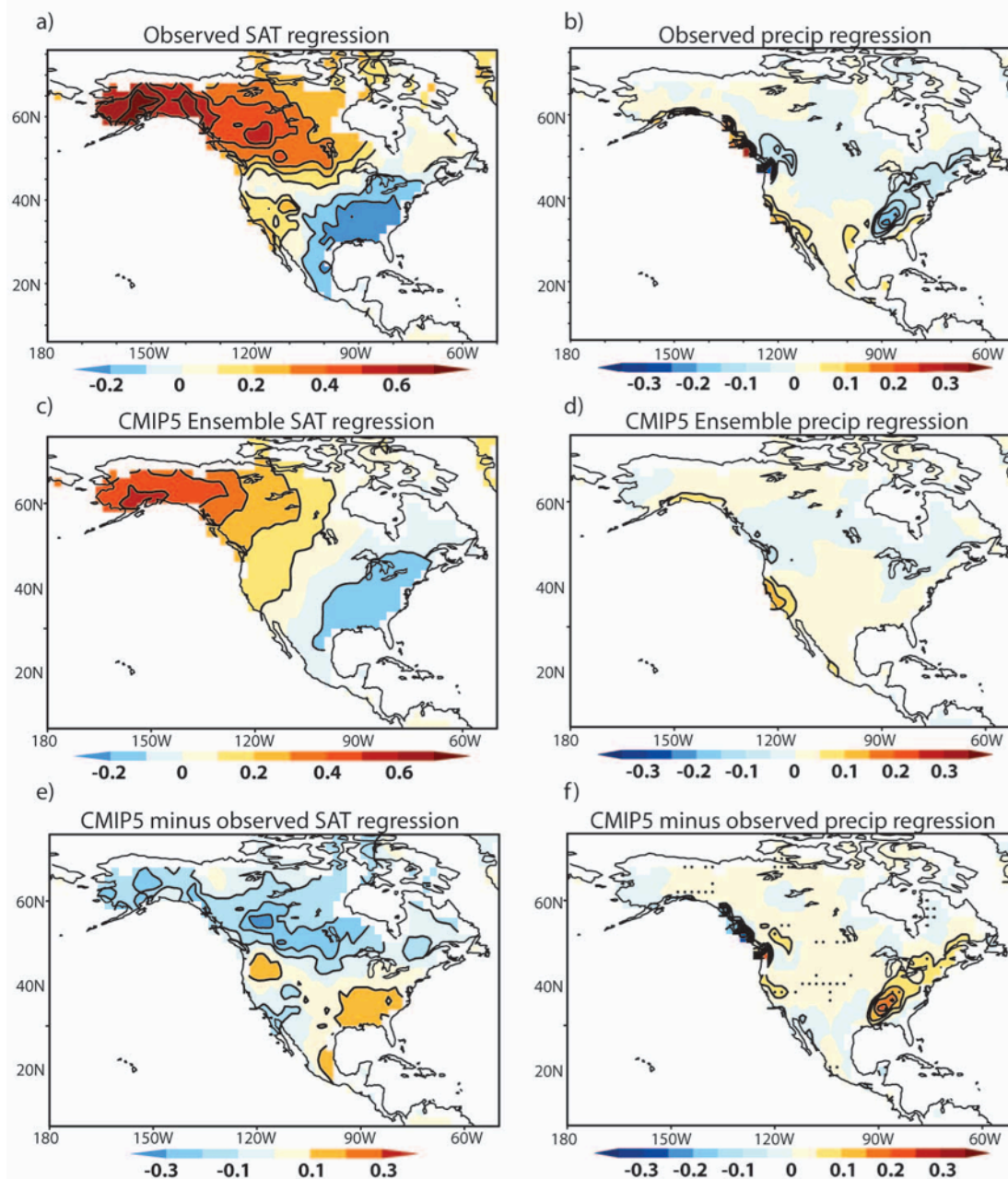


**Figure 16.** Same as Fig. 15 but for CCSM4 simulation results.



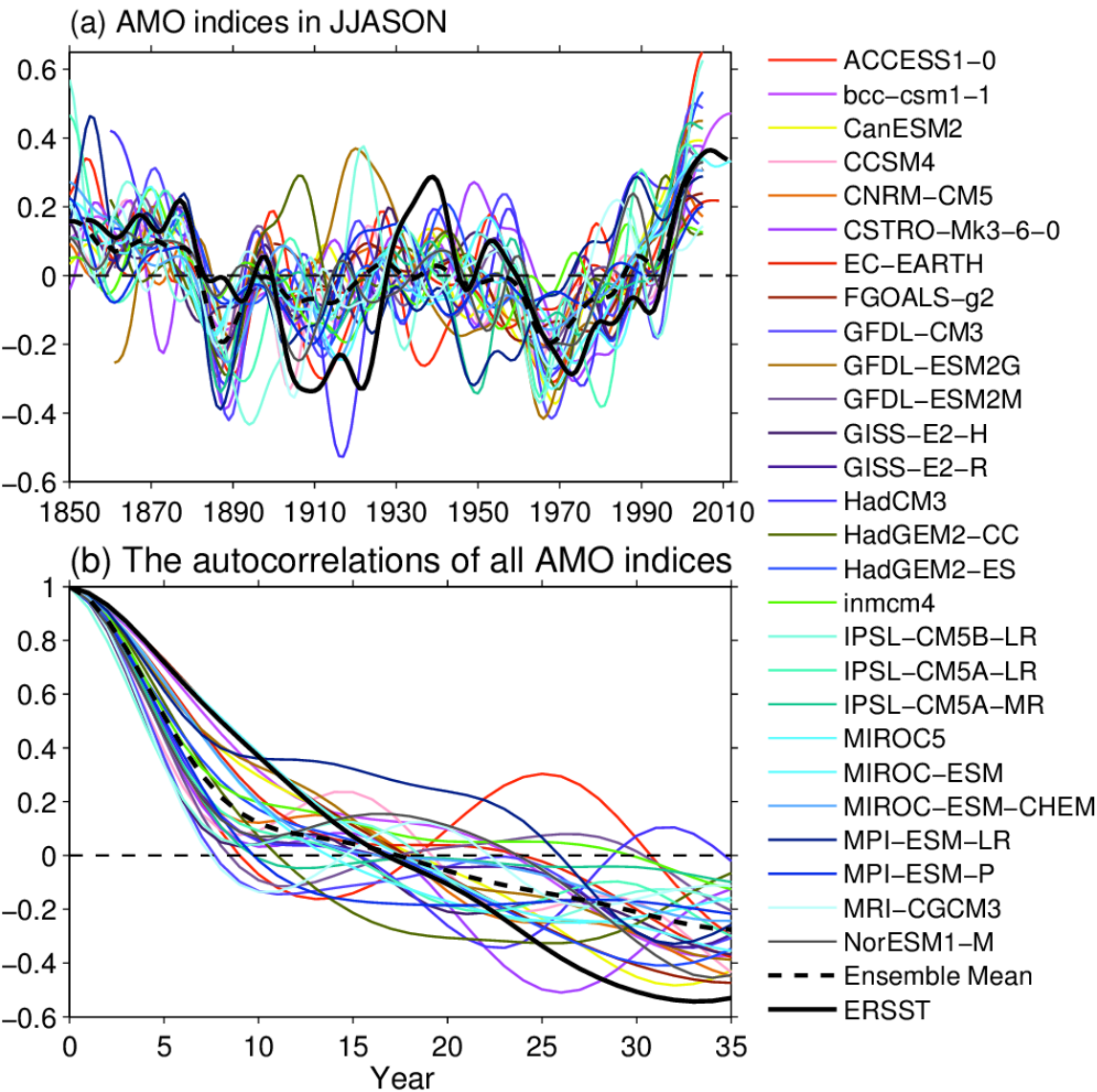


**Figure 17.** PDO SST patterns in observations and CMIP5 models. Linear regression of SST on the PDO index in (a) observations and (b) the CMIP5 ensemble, and (c) the CMIP5 minus observed PDO regression. Observations are from the HadISST dataset for the period between 1870 and 2009. For the CMIP5 models, the analysis period begins as early as 1850 and extends to 2005, and a single realization is used for each model. The contour interval is  $0.2^{\circ}\text{C}$  in (a) and (b) and  $0.1^{\circ}\text{C}$  in (c), with the zero contour omitted. Stippling in (c) indicates where the differences are statistically significant at the 95% significance level based on a two-sided t-test. (d) Standard deviation of the PDO SST regressions within the ensemble. Contour interval is  $0.05^{\circ}\text{C}$ .



**Figure 18.** December-February PDO SAT and precipitation regression patterns over North America. Regressions of DJF SAT (a,c) and precipitation (b,d) on the PDO index in (a,b) observations and (c,d) the CMIP5 ensemble. The differences between the regression patterns (CMIP5 minus observations) are shown in (e) and (f). The contour interval is  $0.1^{\circ}\text{C}$  for the SAT regressions (a, c, e) and  $0.05 \text{ mm/day}$  for the precipitation regressions (b,d,f). Stippling in (e) and (f) correspond to differences that are significantly different at the 95% confidence level based on a two-sided t-test. To focus on multidecadal variability a Butterworth 10-year low-pass filter is applied to each PDO index time series, which is then re-standardized and detrended. The SAT and precipitation anomalies are then regressed on the filtered index for each season. The observations are the CRU TS3.1 temperature and precipitation datasets.

1812



1813

1814 **Figure 19.** The JJASON AMO index in CMIP5 models compared to observations for (a)  
1815 the time series and (b) autocorrelations. The AMO index is defined as the the 11-year  
1816 running mean of the detrended North Atlantic SST during the Atlantic hurricane season  
1817 of June to November (JJASON) from the equator to 60°N, 75°W-5°W. SST observations  
1818 are from the ERSST dataset.

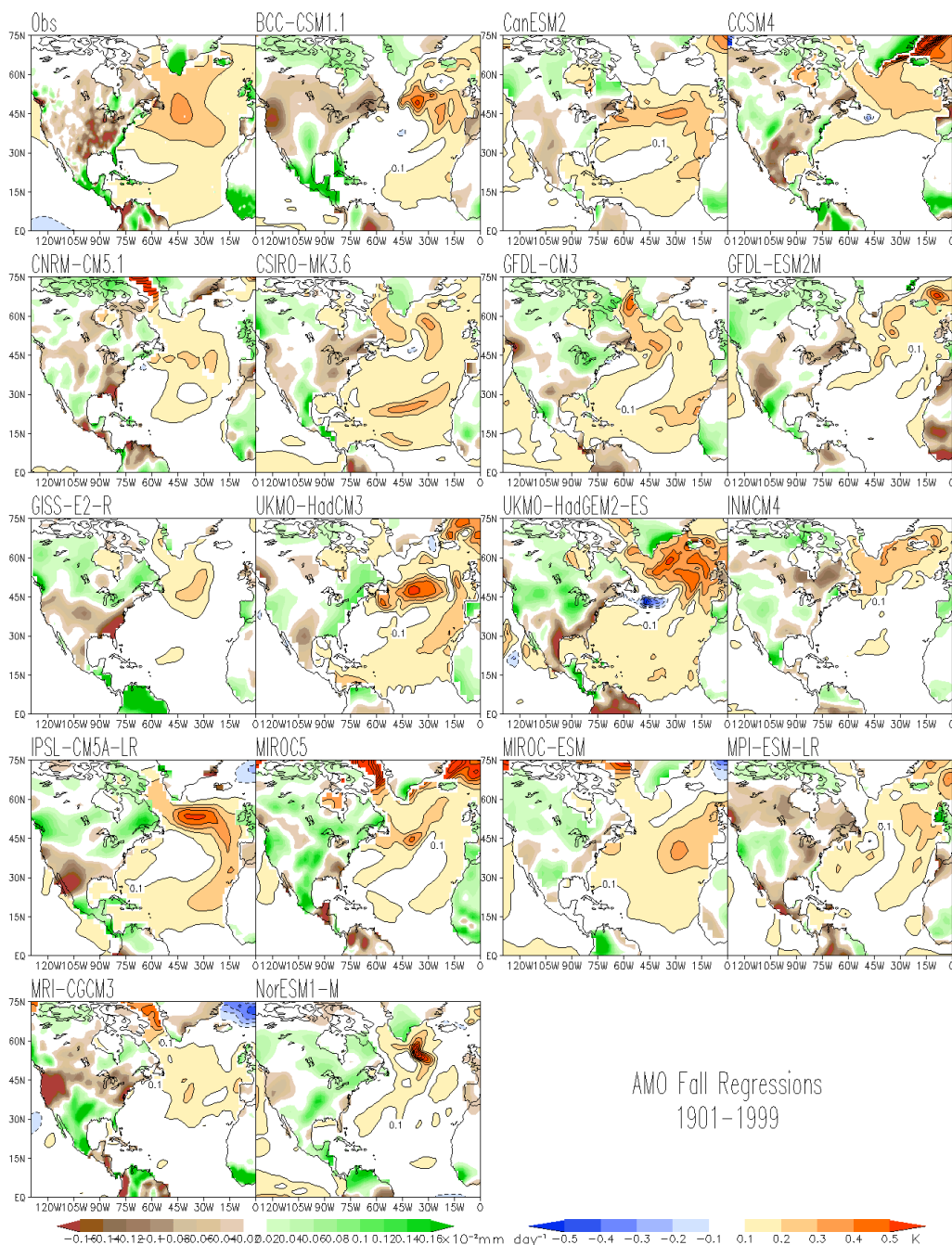
1819

1820

1821

1822



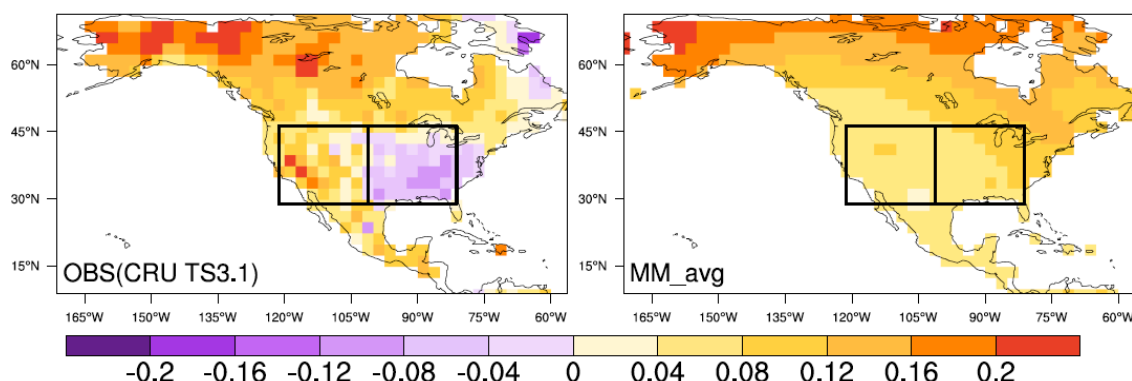


1823

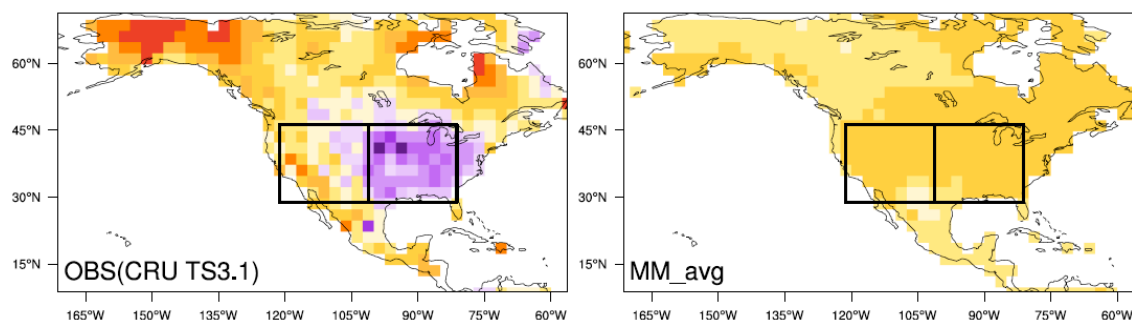
1824 **Figure 20.** Autumn (SON) regressions of the AMO index on SST and precipitation from  
 1825 observations (HadISSTv1.1 and CRU TS3.1), and 17 CMIP5 models for 1901-1999. The  
 1826 AMO index is the area-averaged SST anomalies over the domain (75°-5°W, 0°-60°N),  
 1827 which are detrended and then smoothed via a 11-year running mean. Regressions are  
 1828 calculated for the first ensemble member for each model; observed and simulated  
 1829 anomalies have been regridded to a 1.5°×1.5° grid for precipitation, and a 5°×2.5°  
 1830 grid for SST. Blue/red shading denotes negative/positive SST anomalies, while brown/green  
 1831 shading denotes negative/positive precipitation anomalies. Contour interval is 0.1K and  
 1832 0.02 mm day<sup>-1</sup>, respectively.



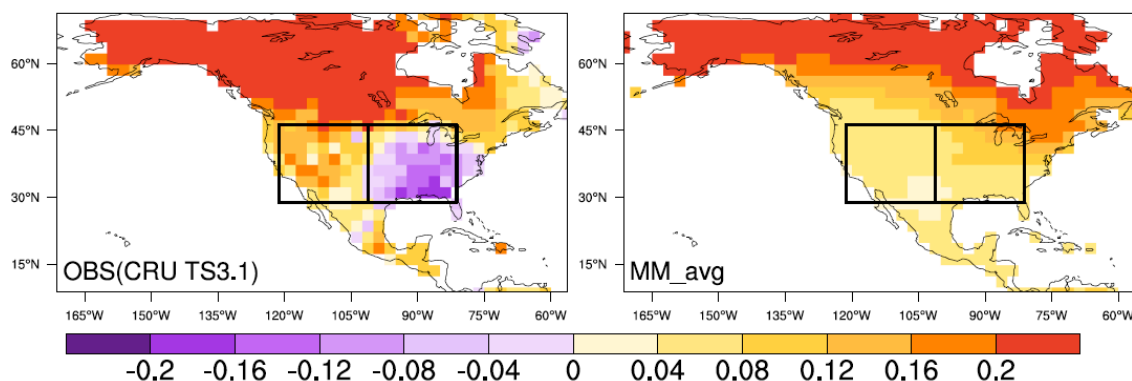
### (a) Annual Temperature Trend ( $^{\circ}\text{C}/\text{decade}$ ), 1930-2004



### (b) Summer (JJA) Temperature Trend ( $^{\circ}\text{C}/\text{decade}$ ), 1930-2004

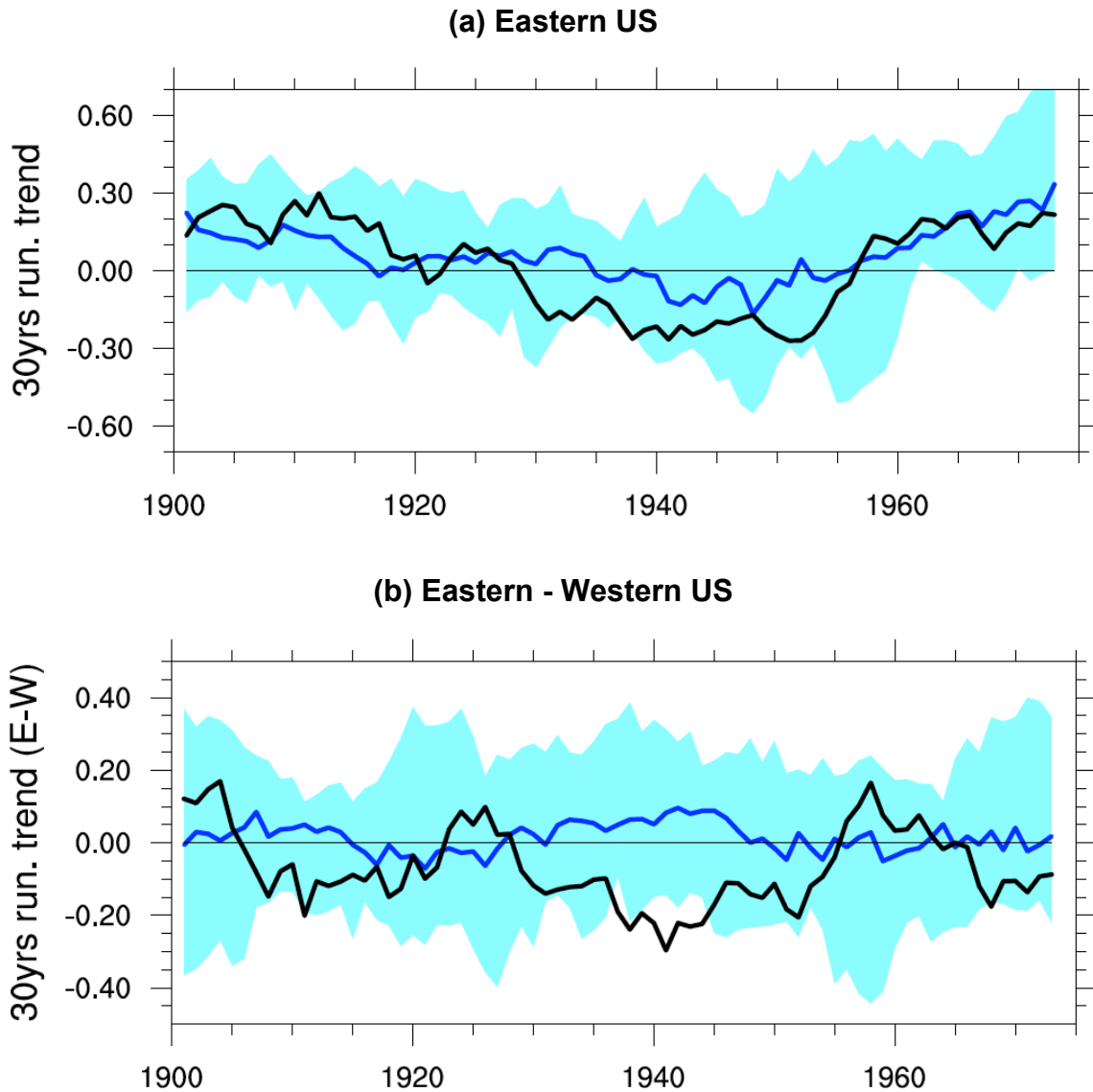


### (c) Winter (DJF) Temperature Trend ( $^{\circ}\text{C}/\text{decade}$ ), 1930-2004



**Figure 21.** Observed and MME mean temperature trends ( $^{\circ}\text{C}/\text{decade}$ ) for North America (1930-2004) for (a) annual, (b) summer, and (c) winter. Observations are from the CRU TS3.1 dataset. The MME mean is calculated from the first ensemble member of 17 models (BCC-CSM1.1, CanESM2, CCSM4, CNRM-CM5, CSIRO-Mk3-6-0, GFDL-CM3, GFDL-ESM2M, GISS-E2-R, HadCM3, HadGEM2-ES, INMCM4, IPSL-CM5A-LR, MIROC5, MIROC-ESM, MPI-ESM-LR, MRI-CGCM3, NorESM1-M). Eastern and western US regions are shown by the boxes.

1840



1841

1842 **Figure 22:** 30-year running annual temperature trend for (a) the Eastern US, (b)  
 1843 difference in trend between the Eastern and Western US. Regions are defined in Fig. 21.  
 1844 Shading represents the 95% uncertainty range calculated from 17 models (see Fig. 21),  
 1845 one ensemble member from each model. Black solid line is the observation (CRU TS3.1)  
 1846 and blue solid line is the MME median. X-axis represents the start of the 30-year running  
 1847 period. For example, the trend value at 1930 represents the trend from 1930 to 1959.

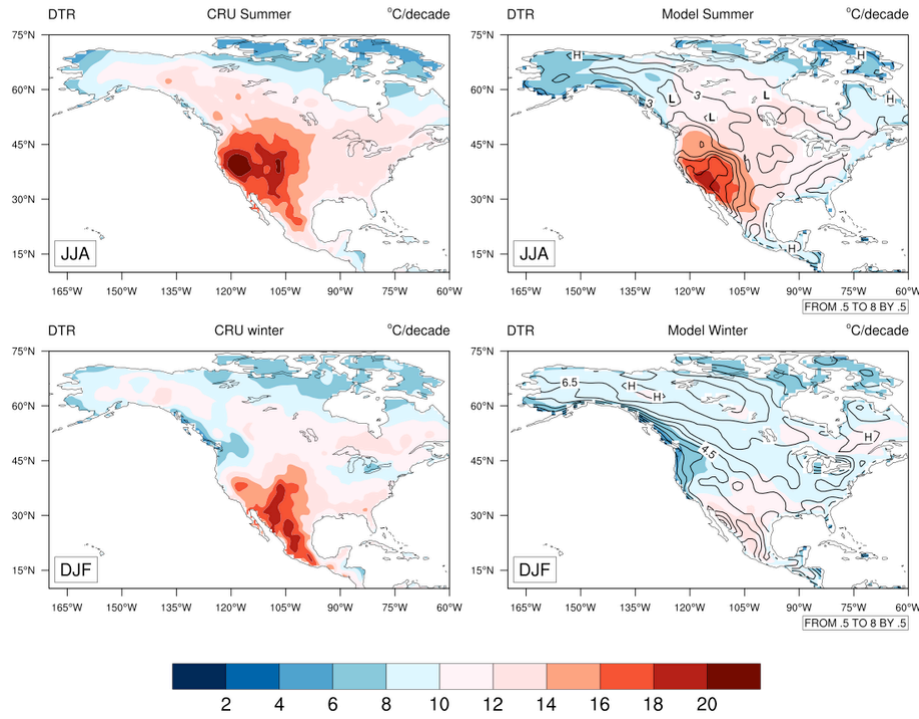
1848

1849

1850

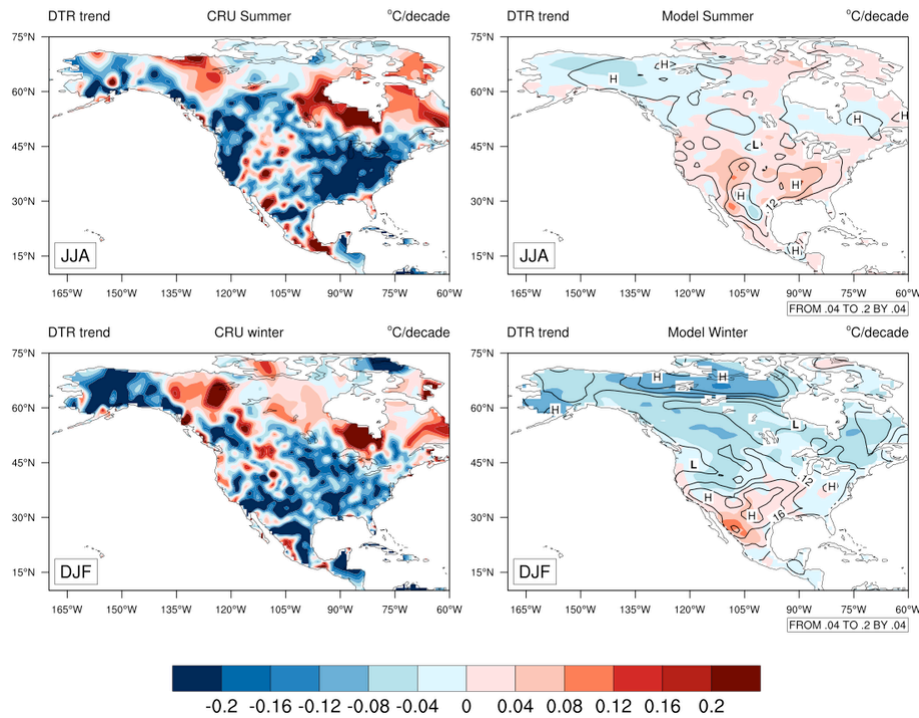
(a)

### Daily Temperature Range (DTR) during 1951-2000



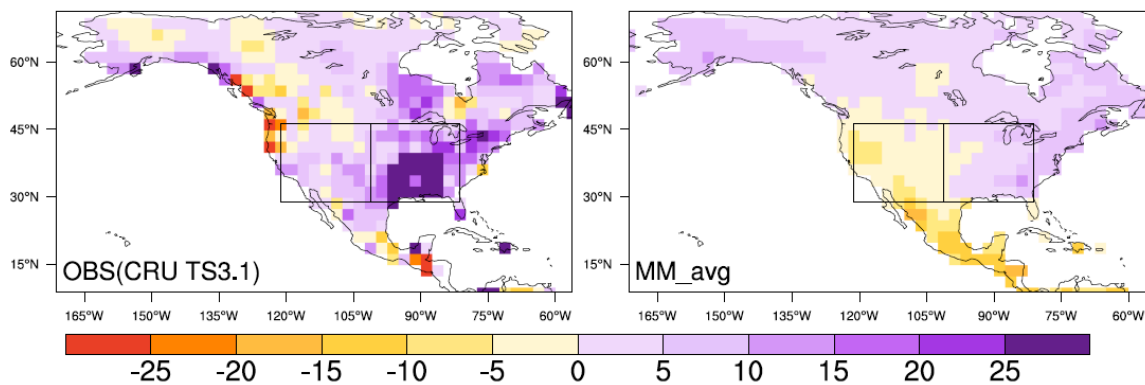
(b)

### Trend of Daily Temperature Range (DTR) during 1951-2000



**Figure 23.** Comparison of (a) mean DTR and (b) DTR trend for the obsevrations (CRU TS3.1) and MME mean of the 17 models (right; see Fig. 21) for 1951-2000. The contour lines are the inter-model standard deviation.

1854



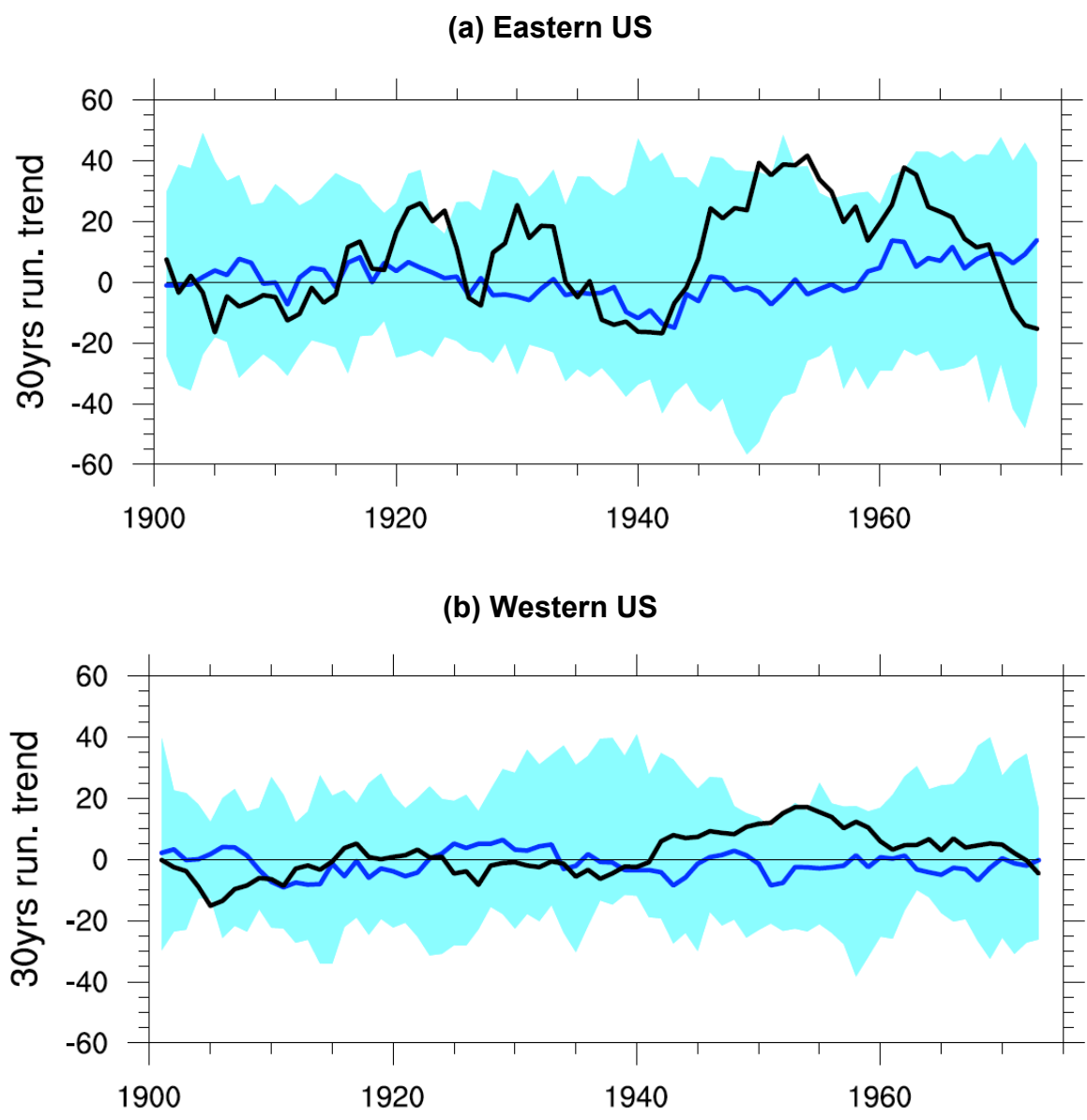
1855

1856 **Figure 24.** Observed and MME mean annual precipitation trends (mm/decade) for North  
1857 America (1930-2004). Observations are from the CRU TS3.1 dataset. The MME mean is  
1858 from 17 models (see Fig. 21), 1 ensemble member from each model. Eastern and western  
1859 US regions are shown by the boxes.

1860

1861

1861



1862

1863 **Figure 25.** 30-year running annual precipitation trend (mm/decade) for (a) the Eastern  
1864 US, and (b) the Western US. Regions are shown in Fig. 23. The shaded region is the 95%  
1865 uncertainty range calculated from 17 models (see Fig. 21), one ensemble member from  
1866 each model. Black solid line is the observation (CRU TS3.1) and blue solid line is the  
1867 MME median. X-axis represents the start of the 30-year running period. For example, the  
1868 trend value at 1930 represents the trend from 1930 to 1959.

1869